

```

1 TCCTCCGGTC GCCCGCCCTC GGGGCAGCTA GTGGCGCAGC CCCCCGCCCCG
51 CGGCCCTGGC CTCCCGGGCG GCGCGGCAGG GGAGGGGTTA AGCTGCCGCA
101 GGGACCGCCG CGTGCGGGGC GAGAGGGAGC CCCCCTGGG GGTGGCGCAG
151 CCGGCGGGGT TCGGTCCGAG CCCGGTGGGA GGCTCCCGGA GCGCAGCCTG
201 GGGCCAGCCC ACCCGCGGCC GCGGGCCATG GCAGGCACCC TGGACCTGGA
251 CAAGGGCTGC ACGGTGGAGG AGCTGCTCCG CGGGTGCATC GAAGCCTTCG
301 ATGACTCCGG GAAGGTGCGG GACCCGCAGC TGGTGCGCAT ATTCTCATG
351 ATGCACCCCT GGTACATCCC CTCTCTCAG CTGGCGGCCA AGCTGCTCCA
401 CATCTACCAA CAATCCCGGA AGGACAATC CAATCCCTG CAGGTGAAAA
451 CGTGCCACCT GGTGAGGTAC TGGATCTCCG CCTTCCAGC GGAGTTTGAC
501 TTGAACCGCG AGTTGGCTGA GCAGATCAAG GAGCTGAAGG CTCTGCTAGA
551 CCAAGAAGGG AACCGACGGC ACAGCAGCCT AATCGACATA GACAGCGTCC
601 CTACCTACAA GTGAAGCGG CAGGTGACTC AGCGGAACCC TGTGGGACAG
651 AAAAAGCGCA AGATGTCCCT GTTGTGTTGAC CACCTGGAGC CCATGGAGCT
701 GCGGGAGCAT CTCACCTACT TGGAGTATCG CTCCTTCTGC AAGATCCTGT
751 TTCAGGACTA TCACAGTTTC GTGACTCATG GCTGCACTGT GGACAACCCC
801 GTCCTGGAGC GGTTCATCTC CCTCTTCAAC AGCGTCTCAC AGTGGGTGCA
851 GCTCATGATC CTCAGCAAAC CCACAGCCCC GCAGCGGGCC CTGGTCATCA
901 CACACTTTGT CCACGTGGCG GAGAAGCTGC TACAGCTGCA GAACCTCAAC
951 ACGCTGATGG CAGTGGTCGG GGGCCTGAGC CACAGCTCCA TCTCCCGCCT
1001 CAAGGAGACC CACAGCCACG TTAGCCCTGA GACCATCAAG CTCTGGGAGG
1051 GTCTCACGGA ACTAGTGACG GCGACAGGCA ACTATGGCAA CTACCGGCGT
1101 CGGCTGGCAG CCTGTGTGGG CTTCCGCTTC CCGATCCTGG GTGTGCACCT
1151 CAAGGACCTG GTGGCCCTGC AGCTGGCACT GCCTGACTGG CTGGACCCAG
1201 CCGGACCCG GCTCAACGGG GCCAAGATGA AGCAGCTCTT TAGCATCCTG
1251 GAGGAGCTGG CCATGGTGAC CAGCCTGCGG CCACCAGTAC AGGCCAACCC
1301 CGACCTGCTG AGCCTGCTCA CCGTGTCTCT GGATCAGTAT CAGACGGAGG
1351 ATGAGCTGTA CCAGCTGTCC CTGCAGCGGG AGCCGCGCTC CAAGTCCTCG
1401 CCAACCAGCC CCACGAGTTG CACCCACCA CCCCAGCCCC CGGTACTGGA
1451 GGAGTGGACC TCGCTGCCA AACCCAAGCT GGATCAGGCC CTCGTGGTGG
1501 AGCACATCGA GAAGATGGTG GAGTCTGTGT TCCGGAACCT TGACGTCGAT
1551 GGGGATGGCC ACATCTCACA GGAAGAATTC CAGATCATCG GTGGGAACCT
1601 CCCTTACCTC AGCGCCTTTG GGGACCTCGA CCAGAACCAG GATGGCTGCA
1651 TCAGCAGGGA GGAGATGGTT TCCTATTTCC TCGCTCCAG CTCTGTGTTG
1701 GGGGGGCGCA TGGGCTTCGT ACACAACCTC CAGGAGAGCA ACTCCTTGCG
1751 CCCCCTCGCC TGCCGCCACT GCAAAGCCCT GATCCTGGGC ATCTACAAGC
1801 AGGGCCTCAA ATGCCGAGCC TGTGGAGTGA ACTGCCACAA GCAGTGCAAG
1851 GATCGCTGT CAGTTGAGTG TCGGCGCAGG GCCCAGAGTG TGAGCCTGGA
1901 GGGGTCTGCA CCTCACCTT CACCCATGCA CAGCCACCAT CACCGCGCCT
1951 TCAGCTTCTC TCTGCCCCG CCTGGCAGGC GAGGCTCCAG GCCTCCAGCA
2001 ATCCCCCTCC CAGCAGAGAT CCGTGAGGAG GAGGTACAGA CGGTGGAGGA
2051 TGGGGTGTTC GACATCCACT TGTAATAGAT GCTGTGGTTG GATCAAGGAC
2101 TCATTCTGTC CTTGGAGAAA ATAATTCAAC CAGAGCAGGG AGCCTGGGGG
2151 TGTGCGGGCA GGAGGCTGGG GATGCGGGTG GGATATGAGG GTGGCATGCA
2201 GCTGAGGGCA GGGCCAGGGC TGGTGTCCCT AAGGTTGTAC AGACTCTTGT
2251 GAATATTTGT ATTTTCCAGA TGAATAAAA AGGCCCGTGT AATTAATAAAA
2301 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA (SEQ ID NO:1)

```

**FEATURES:**

```

5'UTR:      1-227
Start Codon: 228
Stop Codon:  2073
3'UTR:      2076

```

FIGURE 1, page 1 of 2

**Homologous proteins:**Top 10 BLAST Hits

	Score	E
CRA 1000682340958 /altid=gi 6358505 /def=gb AAF07219.1 AF043722...	1293	0.0
CRA 18000005086608 /altid=gi 5031623 /def=ref NP_005816.1  RAS ...	1241	0.0
CRA 18000005188697 /altid=gi 6755290 /def=ref NP_035372.1  RAS,...	1202	0.0
CRA 18000005205935 /altid=gi 7662334 /def=ref NP_056191.1  KIAA...	618	e-175
CRA 18000005188699 /altid=gi 3928857 /def=gb AAC79700.1  (AF081...	533	e-150
CRA 18000005152782 /altid=gi 9507035 /def=ref NP_062084.1  RAS ...	531	e-149
CRA 18000005192860 /altid=gi 7242201 /def=ref NP_035376.1  RAS ...	529	e-149
CRA 18000005192861 /altid=gi 4038292 /def=gb AAC97349.1  (AF106...	526	e-148
CRA 18000005188698 /altid=gi 5032025 /def=ref NP_005730.1  RAS ...	525	e-148
CRA 1000733831533 /altid=gi 6650545 /def=gb AAF21898.1 AF081197...	525	e-148

BLAST dbEST hits:

	Score	E
gi 5432583 /dataset=dbest /taxon=9606 ...	1310	0.0
gi 9876673 /dataset=dbest /taxon=960...	1281	0.0
gi 11286864 /dataset=dbest /taxon=96...	1249	0.0
gi 11285315 /dataset=dbest /taxon=96...	1207	0.0
gi 5432584 /dataset=dbest /taxon=9606 ...	733	0.0
gi 4372300 /dataset=dbest /taxon=9606 ...	720	0.0
gi 12295751 /dataset=dbest /taxon=96...	700	0.0
gi 12288965 /dataset=dbest /taxon=96...	599	e-168
gi 6920402 /dataset=dbest /taxon=960...	573	e-161
gi 2005039 /dataset=dbest /taxon=9606 ...	573	e-161

**EXPRESSION INFORMATION FOR MODULATORY USE:**

library source:

From BLAST dbEST hits:

gi|5432583 Testis  
gi|9876673 Liver-non-cancerous  
gi|11286864 Brain glioblastoma  
gi|11285315 Brain glioblastoma  
gi|5432584 Testis  
gi|4372300 B Cell Chronic lymphatic leukemia  
gi|12295751 Adult marrow  
gi|12288965 Adult marrow  
gi|6920402 Lymph germinal center B cell  
gi|2005039 Lymph

From tissue screening panels:

Leukocyte

```

1 MAGTLDLDKG CTVEELLRGC IEAFDDSGKV RDPQLVRIFL MMHPWYIPSS
51 QLAAKLLHIY QQSRKDNSNS LQVKTCHLVR YWISAFPAEF DLNPELAEQI
101 KELKALLDQE GNRHSSLID IDSVPYTKWK RQVTQRNPVG QKKRKMSLLF
151 DHLEPMELAE HLTLEYRSF CKILFQDYHS FVTHGCTVDN PVLERFISLF
201 NSVSQWVQLM ILSKPTAPQR ALVITHFVHV AEKLLQLQNF NTLMAVVGGL
251 SHSSISRLKE THSHVSPETI KLWEGLTELV TATGNYGNR RRLAACVGFR
301 FPILGVHLKD LVALQLALPD WLDPARTRLN GAKMKQLFSI LEELAMVTSL
351 RPPVQANPDL LSLTIVSLDQ YQTEDELYQL SLQREPRSKS SPTSPTSCTP
401 PPRPPVLEEW TSAAKPKLDQ ALVVEHIEKM VESVFRNFDV DGDGHISQEE
451 FQIIRGNFPY LSAFGDLDQN QDGCISREEM VSYFLRSSSV LGGRMGFVHN
501 FQESNSLRPV ACRHCKALIL GIYKQGLKCR ACGVNCHKQC KDRLSVECCR
551 RAQSVSLEGS APSPPMHS HHRAFSFSLP RPGRGRSRPP AIPLPAEIRE
601 EEVQTVEDGV FDIHL (SEQ ID NO:2)

```

# **FEATURES:**

## **Functional domains and key regions:**

[1] PDOC00004 PS00004 CAMP\_PHOSPHO\_SITE  
cAMP- and cGMP-dependent protein kinase phosphorylation site

Number of matches: 3

```

1      113-116 RRHS
2      144-147 RKMS
3      584-587 RRGs

```

[2] PDOC00005 PS00005 PKC\_PHOSPHO\_SITE  
Protein kinase C phosphorylation site

Number of matches: 7

```

1      27-29 SGK
2      63-65 SRK
3      126-128 TYK
4      134-136 TQR
5      269-271 TIK
6      349-351 SLR
7      506-508 SLR

```

[3] PDOC00006 PS00006 CK2\_PHOSPHO\_SITE  
Casein kinase II phosphorylation site

Number of matches: 9

```

1      12-15 TVEE
2      63-66 SRKD
3      117-120 SLID
4      163-166 TYLE
5      339-342 SILE
6      373-376 TEDE
7      447-450 SQEE
8      476-479 SREE
9      605-608 TVED

```

[4] PDOC00008 PS00008 MYRISTYL  
N-myristoylation site

Number of matches: 4

```

1      19-24 GCIEAF
2      249-254 GLSHSS
3      284-289 GNYGNY
4      492-497 GGRMGF

```

FIGURE 2, page 1 of 7

[5] PDOC00009 PS00009 AMIDATION  
Amidation site

582-585 PGRR

[6] PDOC00018 PS00018 EF\_HAND  
EF-hand calcium-binding domain

Number of matches: 2  
1 439-451 DVDGDGHISQEEF  
2 468-480 DQNQDGCISREEM

[7] PDOC00379 PS00479 DAG\_PE\_BIND\_DOM\_1  
Phorbol esters / diacylglycerol binding domain

499-548 HNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVEC

Membrane spanning structure and domains:

Helix	Begin	End	Score	Certainty
1	34	54	0.713	Putative
2	195	215	0.653	Putative
3	238	258	0.788	Putative

**BLAST Alignment to Top Hit:**

>CRA|18000005086608 /altid=gi|5031623 /def=ref|NP\_005816.1| RAS guanyl  
releasing protein 2 (calcium and DAG-regulated); calcium  
and diacylglycerol-regulated guanine nucleotide exchange  
factor I [Homo sapiens] /org=Homo sapiens /taxon=9606  
/dataset=nraa /length=609

Length = 609

Score = 1241 bits (3176), Expect = 0.0

Identities = 608/615 (98%), Positives = 609/615 (98%)

Frame = +3

Query: 228 MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIY 407  
MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIY  
Sbjct: 1 MAGTLDLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIY 60

Query: 408 QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 587  
QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID  
Sbjct: 61 QQSRKDNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLID 120

Query: 588 IDSVPPTYKWKQRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 767  
IDSVPPTYKWKQRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS  
Sbjct: 121 IDSVPPTYKWKQRQVTQRNPVGQKKRKMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768 FVTHGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 947  
FVTHGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF  
Sbjct: 181 FVTHGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNF 240

Query: 948 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR 1127  
NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR  
Sbjct: 241 NTLMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNRRRLAACVGFR 300

Query: 1128 FPILGVHLKDLVALQALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL 1307  
FPILGVHLKDLVALQALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL  
Sbjct: 301 FPILGVHLKDLVALQALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDL 360

Query: 1308 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 1487  
LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ  
Sbjct: 361 LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 420

Query: 1488 ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 1667  
ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM  
Sbjct: 421 ALVVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNQDGCISREEM 480

Query: 1668 VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847  
VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC  
Sbjct: 481 VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848 KDRLSVECRRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPAIPAEIRE 2027  
KDRLSVECRRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP EIRE  
Sbjct: 541 KDRLSVECRRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP-----EIRE 594

Query: 2028 EEVQTVEDGVFDIHL 2072  
EEVQTVEDGVFDIHL  
Sbjct: 595 EEVQTVEDGVFDIHL 609 (SEQ ID NO:4)

>CRA|1000682340958 /altid=gi|6358505 /def=gb|AAF07219.1|AF043722\_1  
 (AF043722) guanine exchange factor MCG7 isoform 1 [Homo  
 sapiens] /org=Homo sapiens /taxon=9606 /dataset=nraa  
 /length=671  
 Length = 671  
 Score = 1293 bits (3309), Expect = 0.0  
 Identities = 639/670 (95%), Positives = 643/670 (95%), Gaps = 4/670 (0%)  
 Frame = +3

Query: 75 GRGGVKLPQGPPRAGREGAPGGGGAAG----GVRSEPGGRLPERSLGPAHPAPAAMAGTL 242  
 GRG P + +E G +G GVRSEPGGRLPERSLGPAHPAPAAMAGTL  
 Sbjct: 8 GRGTQGWPGSSEQHVQEATSSAGLHSGVDELGVRSEPGGRLPERSLGPAHPAPAAMAGTL 67

Query: 243 DLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRK 422  
 DLDKGCTVEELLRGCIEAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLAAKLLHIYQQSRK  
 Sbjct: 68 DLDKGCTVEELLRGCIEAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLAAKLLHIYQQSRK 127

Query: 423 DNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVP 602  
 DNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVP  
 Sbjct: 128 DNSNSLQVKTCHLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDIDSVP 187

Query: 603 TYKWKQRQVTQRNPVGQKKRMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 782  
 TYKWKQRQVTQRNPVGQKKRMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG  
 Sbjct: 188 TYKWKQRQVTQRNPVGQKKRMSLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFVTHG 247

Query: 783 CTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLQLQNFNTLMA 962  
 CTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLQLQNFNTLMA  
 Sbjct: 248 CTVDNPVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLQLQNFNTLMA 307

Query: 963 VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRFPILG 1142  
 VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRFPILG  
 Sbjct: 308 VVGGLSHSSISRLKETHSHVSPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRFPILG 367

Query: 1143 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT 1322  
 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT  
 Sbjct: 368 VHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLT 427

Query: 1323 VSLDQYQTEDELYQLSLQREPRSKSSPTSPTCTPPPRPPVLEEWTSAAKPKLDQALVVE 1502  
 VSLDQYQTEDELYQLSLQREPRSKSSPTSPTCTPPPRPPVLEEWTSAAKPKLDQALVVE  
 Sbjct: 428 VSLDQYQTEDELYQLSLQREPRSKSSPTSPTCTPPPRPPVLEEWTSAAKPKLDQALVVE 487

Query: 1503 HIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEMVSYFL 1682  
 HIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEMVSYFL  
 Sbjct: 488 HIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEMVSYFL 547

Query: 1683 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS 1862  
 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS  
 Sbjct: 548 RSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLS 607

Query: 1863 VECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPPAIPAEIREEEVQT 2042  
 VECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP EIREEEVQT  
 Sbjct: 608 VECRRRAQSVSLEGSAPSPSPMHSHHHRAFSFSLPRPGRRGSRPP-----EIREEEVQT 661

Query: 2043 VEDGVFDIHL 2072  
 VEDGVFDIHL  
 Sbjct: 662 VEDGVFDIHL 671 (SEQ ID NO:5)

FIGURE 2, page 4 of 7

```

>CRA|18000005188697 /altid=gi|6755290 /def=ref|NP_035372.1| RAS,
    guanyl releasing protein 2; RAP 1A protein-specific
    guanine nucleotide exchange factor 1; CalDAG-GEFI [Mus
    musculus] /org=Mus musculus /taxon=10090 /dataset=nraa
    /length=608
    Length = 608
    Score = 1202 bits (3076), Expect = 0.0
    Identities = 589/615 (95%), Positives = 597/615 (96%)
    Frame = +3

Query: 228  MAGTLDLDKGCTVEELLRGCI EAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLA AKLLHIY 407
           MA TLDLDKGCTVEELLRGCI EAFDDSGKVRDPQLVR+FLMMHPWYIPSSQLA+KLLH Y
Sbjct: 1    MASTLDLDKGCTVEELLRGCI EAFDDSGKVRDPQLVRMFLMMHPWYIPSSQLASKLLHFY 60

Query: 408  QQSRKDNSNSLQVK TCHLVRYWISAFFAEFDLNP ELAEQIKELKALLDQEGNRRHSSLID 587
           QQSRKDNSNSLQVK TCHLVRYW+SAFFAEFDLNP ELAE IKELKALLDQEGNRRHSSLID
Sbjct: 61  QQSRKDNSNSLQVK TCHLVRYWVSAFFAEFDLNP ELAEPIKELKALLDQEGNRRHSSLID 120

Query: 588  IDSVP TYKWK RQVTQRNPV GQKKRKMSLLFDHLEP MELAEHLTYLEYRSFCKILFQDYHS 767
           I+SVPTYKWK RQVTQRNPV QKKRKMSLLFDHLEP MELAEHLTYLEYRSFCKILFQDYHS
Sbjct: 121  IESVP TYKWK RQVTQRNPVEQKKRKMSLLFDHLEP MELAEHLTYLEYRSFCKILFQDYHS 180

Query: 768  FVTHGCTVDNPV LERFISLFNSVSQWVQLMILSKPTAPQ RALVITHFVHVAEKLLQLQNF 947
           FVTHGCTVDNPV LERFISLFNSVSQWVQLMILSKPTA QRALVITHFVHVAEKLLQLQNF
Sbjct: 181  FVTHGCTVDNPV LERFISLFNSVSQWVQLMILSKPTATQ RALVITHFVHVAEKLLQLQNF 240

Query: 948  NTLMAVVGGLSHSSISRLKETHSHVSPETIKLW EGLTELVTATGNYGN YRRRLAACVGFR 1127
           NTLMAVVGGLSHSSISRLKETHSHVSP+TIKLW EGLTELVTATGNY NYRRRLAACVGFR
Sbjct: 241  NTLMAVVGGLSHSSISRLKETHSHVSPDTIKLW EGLTELVTATGNYSNYRRRLAACVGFR 300

Query: 1128  FPILGVHLKDLVALQ LALPDWLDPARTRLNGAKMKQLFSILEELAMVTS LRPPVQANPDL 1307
            FPILGVHLKDLVALQ LALPDWLDP RTRLNGAKM+QLFSILEELAMVTS LRPPVQANPDL
Sbjct: 301  FPILGVHLKDLVALQ LALPDWLDPGRTRLNGAKMRQLFSILEELAMVTS LRPPVQANPDL 360

Query: 1308  LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQ 1487
            LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTS AKPKLDQ
Sbjct: 361  LSLLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSVAKPKLDQ 420

Query: 1488  ALVVEHIEK MVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEM 1667
            ALV EHEK MVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEM
Sbjct: 421  ALVAEHIEK MVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEM 480

Query: 1668  VSYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847
            +SYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC
Sbjct: 481  ISYFLRSSSVLGGRMGFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 540

Query: 1848  KDRLSVECRRRRAQSVSLEGSAPSPSPMHSHHRAFSFSLPRPGRRSRPPAIPAEIRE 2027
            KDRLSVECRRRRAQSVSLEGSAPSPSP H+ HHRAFSFSLPRPGRR SRPP EIRE
Sbjct: 541  KDRLSVECRRRRAQSVSLEGSAPSPSPHTT-HHRAFSFSLPRPGRRSSRPP-----EIRE 593

Query: 2028  EEVQTVEDGVFDIHL 2072
            EEVQTVEDGVFDIHL
Sbjct: 594  EEVQTVEDGVFDIHL 608 (SEQ ID NO:6)

```

FIGURE 2, page 5 of 7

>CRA|18000005205935 /altid=gi|7662334 /def=ref|NP\_056191.1| KIAA0846  
 protein [Homo sapiens] /org=Homo sapiens /taxon=9606  
 /dataset=nraa /length=689  
 Length = 689  
 Score = 618 bits (1576), Expect = e-175  
 Identities = 314/597 (52%), Positives = 409/597 (67%), Gaps = 5/597 (0%)  
 Frame = +3

Query: 234 GTLDLDKGCTVEELLRGCI EAFDDSGKVRDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQ 413  
 G+ L K T++ELL CIE FDD+G++ + L RI L+MH WY+ S++LA KLL +Y+  
 Sbjct: 2 GSSGLGKAATLDELLCTCIEMFDDNGELDNSYLPRIVLLMHRWYLSSTELAEKLLCMYRN 61

Query: 414 SRKDNSNSLQVKTC HLVRYWISAFPAEFDLNPELAEQIKELKALLDQEGNRRHSSLIDID 593  
 + ++ N ++K C+ +RYWI FPAEF+L+ L +E + + Q G +H SLIDI  
 Sbjct: 62 ATGESCEFRKIKICYFMRYWILKFPAEFNLDLGLIRMTEEFREVASQLGYEKHVSLLIDIS 121

Query: 594 SVPTYKWK RQVTQRNPVGQKKRKM SLLFDHLEPMELAEHLTYLEYRSFCKILFQDYHSFV 773  
 S+P+Y W R+VTQR V KK K L LFDHLEP+ELAEHLT+LE++SF +I F DY S+V  
 Sbjct: 122 SIPSYDWMRRVTQRKKVS-KKGKACLLFDHLEPIELAEHLTFLEHKSFRIRISFTDYQSYV 180

Query: 774 THGCTVDNPNVLERFISLFNSVSQWVQLMILSKPTAPQRALVITHFVHVAEKLLQLQNFNT 953  
 HGC +NP LER I+LFN +S+WWQLM+LSKPT QRA VIT F++VA+KLLQL+NFNT  
 Sbjct: 181 IHGCLENNPTLERSIALFNGISKWVQLMVLSKPTPQQRAEVITKFINVAKLLQLKNFNT 240

Query: 954 LMAVVGGLSHSSISRLKETHSHVSPETIKLWEGLTTELVTATGNYGNRRRLAACVGFRFP 1133  
 LMAVVGGLSHSSISRLKETHSH+S E K W +TELV++ GNY NYR+ A C GF+ P  
 Sbjct: 241 LMAVVGGLSHSSISRLKETHSHLSSEVTKNWNEMTELVSSNGNYCNRYKAFADCDGFKIP 300

Query: 1134 ILGVHLKDLVALQLALPDWLDPARTRLNGAKMKQLFSILEELAMVTSLRPPVQANPDLLS 1313  
 ILGVHLKDL+A+ + PDW + ++N KM QL L EL + + ++ N DL++  
 Sbjct: 301 ILGVHLKDLIAVHVIFPDWTE--ENKNVIVKMHQLSVTLSELVSLQASHHLEPNMDLIN 358

Query: 1314 LLTVSLDQYQTEDELYQLSLQREPRSKSSPTSPTSCTPPPRPPVLEEWTSAAKPKLDQAL 1493  
 LLT+SLD Y TED++Y+LSL EPR+ SPTSPT+ P +P V EW PK D +  
 Sbjct: 359 LLTSLDLHYHTEDDIYKLSLVLEPRNSKSPTSPTT---PNKPVVPLEWALGVMPKPDPTV 415

Query: 1494 VVEHIEKMVESVFRNFDVDGDGHISQEEFQIIRGNFPYLSAFGDLDQNDGCISREEMVS 1673  
 + +HI K+VESVFRN+D D DG+ISQE+F+ I NFP+L +F LD++QDG IS++EM++  
 Sbjct: 416 INKHIRKLVESVFRNYDHDGDYISQEDFESIAANFPFLDSFCVLDKQDGLISKDEMMA 475

Query: 1674 YFLRSSSVLGGRM--GFVHNFQESNSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQC 1847  
 YFLR+ S L +M GF+HNFQE L+P C HC + GI KQG KC+ CG NCHKQC  
 Sbjct: 476 YFLRAKSQLHCKMGPGFIHNFQEMTYLKPTFCEHCAGFLWGIKQGYKCKDCGANCHKQC 535

Query: 1848 KDRLSVECRRRRAQSVSL---EGSAPSPSPMHS HHRAFSFSLPRPGRGSRPPAIP 2009  
 KD L + CRR A++ SL GS P + F F G R AI L  
 Sbjct: 536 KDLLVLACRRFARAPSLSSGHGSLPGSPSLPPAQDEVFEFFPGVTAGHRDLDSRAITL 592 (SEQ  
 ID NO:7)

>CRA|18000005188699 /altid=gi|3928857 /def=gb|AAC79700.1| (AF081196)  
 calcium and DAG-regulated guanine nucleotide exchange  
 factor II [Rattus norvegicus] /org=Rattus norvegicus  
 /taxon=10116 /dataset=nraa /length=795  
 Length = 795  
 Score = 533 bits (1358), Expect = e-150  
 Identities = 267/590 (45%), Positives = 390/590 (65%), Gaps = 12/590 (2%)  
 Frame = +3

Query: 156 GVRSEPGGRLPERSLGAHPAPAAMAGTLD-----LDKGCTVEELLRGCI EAFDD 308  
 G R+ P GRL +S PA ++A L KG ++++L+ CI++FD  
 Sbjct: 17 GSRAGPKGRLEAKSTNSPLPAQPSLAQITQFRMMVSLGHLAKGASLDDLIDSCIQSFDAD 76

FIGURE 2, page 6 of 7



Query: 309 GKV-RDPQLVRIFLMMHPWYIPSSQLAAKLLHIYQQSRKDNSNSLQVKTCHLVRYWISAF 485  
 G + R QL+++ L MH I S++L KL+++Y+ + + NS + +K C+ VRYWI+ F  
 Sbjct: 77 GNLCRSNQLLQVMLTMHRIIISAEALLQKLMNLYKDALEKNSPGICLKICYFVRYWITEF 136

Query: 486 PAEFDLNPELAEQIKELKALLDQEGNRRHSSLLIDDSVPTYKWKQVQTQRNPVG-QKKRK 662  
 F ++ L ++E + L+ G H LID + + W R++TQR KKRK  
 Sbjct: 137 WIMFKMDASLTSTMEEFQDLVKANGEESHCHLIDTTQINSRDWSRKLQRIKSNSTSKKRK 196

Query: 663 MSLLFDHLEPMELEHLTYLEYRSFCKILFQDYHSFVTHGCTVDNPVLERFISLFSVSQ 842  
 +SLLFDHLEP EL+EHLTYLE++SF +I F DY +++ + C +NP +ER I+L N +SQ  
 Sbjct: 197 VSLLFDHLEPEELSEHLTYLEFKSFRRISFSDYQNYLVNSCVKENPTMERSIALCNGISQ 256

Query: 843 WVQLMILSKPTAPQRALVITHFVHVAEKLQQLQNFNTLMAVVGGLSHSSISRLKETHSHV 1022  
 WVQLM+LS+PT RA V F+HVA+KL QLQNFNTLMAV+GGL HSSISRLKET SHV  
 Sbjct: 257 WVQLMVLRSRTPQLRAEVFIKFIHVAQKLHQLQNFNTLMAVIGGLCHSSISRLKETSSHV 316

Query: 1023 SPETIKLWEGLTELVTATGNYGNYRRRLAACVGFRRFPILGVHLKDLVALQLALPDWLDPA 1202  
 E K+ +TEL+++ NY NYRR C F+ PILGVHLKDL++L A+PD+L+  
 Sbjct: 317 PHEINKVLGEMTELLSSCRNYDNYRRAYGECTHFKIPILGVHLKDLISLYEAMPDYLEDG 376

Query: 1203 RTRINGAKMKQLFSILEELAMVTSLRPPVQANPDLLSLLTVSLDQYQTEDELYQLSLQRE 1382  
 ++N K+ L++ + EL + + PP+ AN DL+ LLT+SLD Y TEDE+Y+LS RE  
 Sbjct: 377 --KVNQKLLALYNHINELVQLQDVAPPLDANKDLVHLLTSLDLYYTEDEIYELSYARE 434

Query: 1383 PRSKSSPTSPTSCTPPPPPPVLEEWTSAAPKPKLDQALVVEHIEKMVESVFRNFDVDGDGH 1562  
 PR+ +P P +PPV+ +W S PK D + +H+++MV+SVF+N+D+D DG+  
 Sbjct: 435 PRNHRAPP-----LTPSKPPVVVDWASGVSPKPDPKTISKHVQRMVDSVFKNYDLDDQDGY 489

Query: 1563 ISQEEFQIIRGNFPYLSAFGDLQNDGDCISREEMVSYFLRSSSVLGG-RMGFVHNFQES 1739  
 ISQEEF+ I +FP+ +F +D++++G ISR+E+ +YF+R+SS+ +GF HNFQE+  
 Sbjct: 490 ISQEEFEKIAASFPP--SFCVMDKREGLISRDEITAYFMRASSIYSKLGLGFPHNFQET 547

Query: 1740 NSLRPVACRHCKALILGIYKQGLKCRACGVNCHKQCKDRLSVECRRAQS 1889  
 L+P C +C + G+ KQG +C+ CG+NCHKQCKD + EC++R++S  
 Sbjct: 548 TYLKPTFCDCNAGFLWGVKQGYRCKDCGMNCHKQCKDLVVFECKKRSKS 597 (SEQ ID  
 NO:8)

#### Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF00617	RasGEF domain	123.5	4e-33	1
PF00130	Phorbol esters/diacylglycerol binding domain	59.5	3.6e-14	1
PF00036	EF hand	21.8	0.00027	2
PF01237	Oxysterol-binding protein	3.5	4.2	1

#### Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF01237	1/1	249	272 ..	1	24 [.]	3.5	4.2
PF00617	1/1	148	336 ..	1	227 [.]	123.5	4e-33
PF00036	1/2	430	458 ..	1	29 [.]	17.4	0.0047
PF00036	2/2	463	482 ..	5	24 ..	6.7	4.9
PF00130	1/1	499	548 ..	1	51 [.]	59.5	3.6e-14

1 ACAGAAAGGT CCTGTTTCTA AGTCTTACAT TACCAAGACT GAGGTGCGGG  
51 GGCGGTCTTG GATCCCCCGC CCCAAGGCTG GGAGGGGCAC GCCTCGGAAG  
101 GGAGGTTTGG GGTGCGGTGGT TTCACAGTGA GTGTGTCTGA AGCCAAATGG  
151 TCGGAAACCG TTACCCGCTC TCCTAGGCCC GGCTAGTGGG GACCCCAACC  
201 GCCTGCGGCT GCCCTCCCA AGTTCCCTCC TGTGGCCAG GCATCCAGGT  
251 CTCCAGTCTC CGAGCTGCGG AGAACCACCC GCCACATGCG GCTGCCCCCT  
301 TCCATTCGAC CCTGTGGGGA GCCAGGCTTC CGGGGCCCCG TTCCTCCTGT  
351 GTGAACTGGG CCCCCCGCCC CCATTCCAG ACATCAAGGC CGCGTCTCCA  
401 GATAGCCACG ATTTTCATTCC TCGTCCCCA CAGGTCCCTC TCCCCAAAAT  
451 ATTCCCATCT TGTCTAGCC CATCCCCAG ACTATCTCAA GGACCAGCTG  
501 TCCCCACGCC CCCGACCTCC ACTAGGCTG TGCCACCCGC TGCCTGCAGG  
551 AAGACGCCCC GTCCCCGGCC GGGTTAGCCC CATGGGAACG GTTTGTCTCG  
601 AAAACAGGAA CCCGAGCTGG GGGCTGGGCG GGGCGCCCCCT TCCCCACCGC  
651 AGTCCGCTTC CTGCCCCCTC CGGCTTCCTC CGCCCGACAC CCAGGCAGGG  
701 CGGGGGGCAC TGGGGCGTCC GCGGTTGGGG GAGGGGCTCT TCGTTTCGGT  
751 CCCCCCTCCC GCGTCCCGGG CGGCGGGGCC TCCGGTCGCC CGCCTCGGGG  
801 CAGCTAGTGG CGCAGCCCCC CGCCCGCGGC CCTGGCCTCC CGGGCGGCGC  
851 GGCAGGGGAG GGGTTAAGCT GCCGCAGGGA CCGCCGCGTG CGGGGCGAGA  
901 GGGAGCCCCC GGTGGGGGTG GCGCAGCCGG CGGTGCGGAG CTCCGCGCAG  
951 GGGCGAGGGG GGGAGGGGGC AGCCTGGCGC GGGGCGGGGG GCGGGGCGGC  
1001 GGGGAGCGGG GCCGCGCGGT GGAGAGCGGG CGGGAGCCGC AGCCGCAGCG  
1051 AGGCCGCGCG GCGGGAGCGC ACGGAGGTGG GGTGCGCCAG GCCGGTGCGG  
1101 GCTCCTTGCG GCAGGTCCCA AGAGTGAGTG GGCAGCGCGG GCGGGGCGC  
1151 CAGGCGAAGG AGGGCGCGGC CCCAGCGAC TCCCCCCCCG CCCAGGGCGG  
1201 CGCGGGCGGG CTGGGGGCGG CGAGCGGGTG GGGAGTCTGC GGCCCGGGT  
1251 TGGGAGAGGG GGCAGCGGCC ACGAGAGCTA AGGCGCGCTG GATCCCCGGA  
1301 GGGCGGAGGA CCTCCACGGT GCACCCAGCT TTTCCAGACC ACCTTCCAGC  
1351 GGGGCCCTCC CCCGCGTACC CCCATTGGC AGATGAGAAA ATTGAGGCTC  
1401 CCAGAGGCCA AGTGATTCTC AAGGTCACAC GAGGAAGCGG TAGAGCCAGG  
1451 CGGGGACGGC TCTGGGTGGC TCTTAGGAAA AGTCCGCTG AGAACTCCGT  
1501 ACAGGAGCTC CCCTGTCCTC CAGCCTGGGG GAGTGAGTAT GTGTAGGGCC  
1551 GGGGTACCTT TCCGTGGGGC AAGGCTCTGC CAAAATCTGG GAGTGAGGGG  
1601 AGTCAGGGAG CTGGGCGCGC AGGGCGGGCC CTGCACCCGA AATGGGAGGG  
1651 GGGCGACGGA ATGGGCGTGC GCACCCATGG GGGTGTGTGC ATGTGTGTGG  
1701 GAGTGTACAT GCGTGAGAG GCACTGCCTT GCGTGTGTGC ACACGTGTGA  
1751 GGATGTACAG GCCTGTGTGG CCGCGGGACT CAAGGCTGGC CTGGCTCAAG  
1801 TGAACAGCAC GTCCAGGAGG CGACCTCGTC CGCGGGTTG CATTCTGGGG  
1851 TGGACGAGCT GGGTATGTGT GCCTGAGGGT TTCTTCGTGC AGGTGTGCAC  
1901 AGGGTGTGGG TGCCATTGTG TGTGAGAGAC GGAGGATGGG GAGGCGGGTG  
1951 CCTGTGGCCC GGTGCGTGTG AGTGCGGACG CCTGCACCTC CACTTAGGTC  
2001 CCCGCCCTCC GACGACTAAC TTGGGTGTGG AGTGTTTGCC CCTGCCAGGG  
2051 TGCGTATGAC CCCGCCAGTG ACCGGAGTTG CTAATGGTGT CATGCACCCA  
2101 CCGGCCACCC TTGGCGCGAG CGCCCCCTC TGGACACCCT GCTCCGTGCG  
2151 CGCTCACAGT TCGCTGTGCG GGGGCCGGGG CCAGGGTCAG GAGCCGGGGA  
2201 TAGGGAGGAA GAGGCCTGTG GGACAAGCTG AGCCGGGACC CTGGGACCT  
2251 TTGCGGAGGT GGCTGGGAG CGCTCAGTTC CCAGGCTGAG GCTTCCCGCT  
2301 GACGCTCCTT GGCAGCAGCG GGCTCCCCC GCCCAGGAA TGTTCCTCTC  
2351 CCATCCAGTC CGCTCCCTT AGGGCAGGCC CCCTGGGGGC TGCCGCAGCC  
2401 CCGCCTCGCC TTCTGGGCT CCCGGGAGGG GCGGAGGCGA GCAGGACGCC  
2451 TGGGTTCCTT CCCCCACCT CCCATACCAG GGAGAAATTC CTCCGAGGTC  
2501 CCCTCAGGCT CTGGGTTCCC AAAATAACCC TGCGGGGGAA GGGAGGCTGT  
2551 GGAGGGAGGG AAGCGGGAGG GGCGCAGAGC CGAGCTGCGG GGTGCTGCAG  
2601 GTGCTCTTGG GGAGAGGGCG CGAGGAGAAG GCGCCCTGCG GGGGGCTGGG  
2651 CGCCAGCCAG TCCTGGGATC TTGGTTCTGTC CCCATCCTCG TGAAGCCCCCT  
2701 CGGCCTTCCC GCGACTCCGA GGGTGGGCGG GAAGCTCTC TCGGGTCCG  
2751 TTTCCCAACT GGGGGGTTGC ACCATCCCGG GCCAGACCGT TTAACCCCGG  
2801 GAGTGGCCGC GGGGACAAAC TCCGCCCCG TCCAGCAGGG GCGGTGCCCG  
2851 CCCC GCCCG TTTCTGCCCG CGGGGCCGCT CCCCCGCCG CCACTCCGCA  
2901 GACTCCCGCT CTGCTCTCC CGGGACAGGG GTTCGGTCCG AGCCCGGTGG  
2951 GAGGCTCCCG GAGCGCAGCC TGGGCCAGC CCACCCGCG CCGCGGCCA  
3001 TGGCAGGCAC CCTGGACCTG GACAAGGGCT GCACGGTGA GGAGCTGCTC  
3051 CGCGGGTGCA TCGAAGCCTT CCGTGAGTGG CTCGGGAGGG CACACGGAGC  
3101 CTGAGCTAG CCCCAGTCT GAGCCCGGGT CCCTGCCTCC CAGGCACAGT

FIGURE 3, page 1 of 12

3151 CCAGGGCACA GCCCTGACCC GGACCCACCC TGCTCCGCAG CGTGCACTCT  
 3201 CTTTAAACGAA AGCCTCCTCC GCAACGCAGG GCAGAGAGAT GCACGCCCTT  
 3251 CAGACAGATG AGTTTTCCT TCTCTAGCCT TCCCCAGCGG CGGCGAAGGG  
 3301 AGGGCCGGGT CCCGGACTCT GACACTTGAG GGGCATTATC TGTCTCCCGG  
 3351 GGAATCCGGA GGAACCTCGCT ATCTCCGGCC TGGGAGCTGT TTCCGGCTAA  
 3401 TGGGGGGCGG CTTATCTGGT GAAGGGGTGC CCCTTCCCC CAAGCGCTCA  
 3451 GGAAATGACC TCTGGATTCT TGACCCCGGG GAACCCAGGC TCCTTCCGCC  
 3501 CCAGCTGGTT CCCCTCCGGA CGATGGGCGG CTCGGGCGCT CCCCTCCTCC  
 3551 AGTCCTCAGG GCGTGCCTAT CTCTGCCCCA CCACACCTTT CCTCTCTAAT  
 3601 TTGCCTCCTG CTCTCGGAGT CCTGGGCAAG CAGGAGGTGG GCGGGGTGCA  
 3651 GCGTGACCCC GAAGGACCGA TACCTGGCGG GTTGCGGGGT GAGGATGAGG  
 3701 CATGGTAGCT GCGGACCCAG CTCAGCCACC TGTCTTTGAC CCTTCGGAGT  
 3751 CAGATGACTC CGGGAAGGTG CGGGACCCGC AGCTGGTGCG CATGTTCTCTC  
 3801 ATGATGCACC CCTGGTACAT CCCCTCCTCT CAGCTGGCGG CCAAGCTGCT  
 3851 CCACATATAT CTTTCGCCGG CCTTGCCAAG GCGCCCGCCG TCGGAGCCCA  
 3901 TGCGCAGCCC CTCTGCCCAG CCCAGGTGCA GAATGAGCCT CGCTCCTAAG  
 3951 TATAGGCCAC TCCTTATCCC AGAGCTCAGG CGTCGTCCCA GCCTCCAAC  
 4001 AGGGCCTAGG CTCTGCCCCC TCCTTGCTCC TAGCGACTCG GTCCTGTCCC  
 4051 CAGGCTCTGT CCCAGCCGA GGCCTTGCCC TCCTTCTCCC TAGAGTCTAG  
 4101 GGCCTGCCCC TGCTCAGGC TTGGGTGCGC CCCGTGCATC TCTCTCTCCC  
 4151 AGAGCCCAGG CTTTGCTTTC AGCCTCCCTC AGCACCTAGT CCTCCACCCC  
 4201 CACCTCCAAC CCCTCCCAGA GCTCAAGCCT CACCCCAAGC ATCTCCGCAG  
 4251 AGCGCAAGCC CCATCCCTAG AACGTGTCTC CTAGAACCAG GCGCCGCCCC  
 4301 CAGCCTCCCT CCACGCAGGC CTCCCTTTCT AGAGTTAAGC GGCCTCCTTA  
 4351 ACCCTCTCCT TCACCTACCA ACAATCCCGG AAGGACAAC CCAATTCCTT  
 4401 GCAGGTGAAA ACGTGCCACC TGGTCAGGTG AGTCTTTCCC CTGGGGCTCT  
 4451 AGCCCCCTCC CTTTCTCCCT TCTCTCTGGC TTCAGGCTGG CCTGGAGGAG  
 4501 GGGGCAGGGC GCTGTTTCTG GGAGTGGGTG TGAACCCTGG CTGTGCCGGG  
 4551 TGGGCAGTGC TGCCACAGGC TCACCCCTTC CTGGGTCTGG GCCTTAATTT  
 4601 TCTTTTCTGC GCAGTGCGGG TGGTTGTCTC AAGGTCTAA TGTACACTTG  
 4651 GAGTGCGCAA GGAAAGAGCT GGAACCATAG TTTGAGGGTC TTTTGTCTTA  
 4701 GGTGACTATA ATCTCAAATA GTCCTTGCA GCCTGCTGGG TGATGGTGGG  
 4751 GGAAGGGCTA TCTTGGGTGA CTCCCGCTC CTCCAGGTAC TGGATCTCCG  
 4801 CCTTCCCAGC GGAGTTTGAC TTGAACCCGG AGTTGGCTGA GCAGATCAAG  
 4851 GAGCTGAAGG CTCTGCTAGA CCAAGAAGGG AACCAGCGC ACAGCAGCCT  
 4901 AATCGACATA GACAGCGTGT GCGTGGGGGG AGCAGAGAGG GCTGGGGGGG  
 4951 CACTCAGTAT CCTATACCAT CTGTGCTTAA TAAATGTCTG TTGAACTGAA  
 5001 TGAGTGAGGG TCATGTTGCT CTCTCGCTTA AAAACCTTCC ATGGCTCCCT  
 5051 ATTGCCTTCA ACATGCCTCC TCTGGGCAGC TTGGCGTTCC TGCCTCATCT  
 5101 TCCACTGCCA CCACCCATCC CACACACCTC CTCCTGTAGC TGCCTGGGT  
 5151 CGGCTCCCCG TCGGCTGAGC TCTCGAGTCC TTTCTCATCA TGGTGTCTG  
 5201 CTCATATCAT CCCCTTGCT GCCTCCTCCG TGTTACCAAG ACTCAGTTCA  
 5251 GGCATGAAGT CTCGTGGGGC TCTGAGGGT CGGGGCTCTT CCGGGGTAGA  
 5301 ATTTGTCTGTT CCCACCTCTG TTTTCCATGG CACTTTGTAC AGACTCCTGT  
 5351 ACAAGACCT CTGTACATGT GTCACGCTGT TTTGTGATCA TGTGTTTCTG  
 5401 TGTCTGTCTC CCTCAGTAGA CTGTGAGCTC CTCGAGGGCA GGAACCGTGT  
 5451 CTTACTCATC TCTGTATTCC CAGCGCCTAG CACAGTGCCT GGCACAGAGT  
 5501 ACGTTGTTCA TAAATGTGTG TTGAGTGCAT GACGGGGTGG GGGGAGATGA  
 5551 GGAGGAGTTG CTGGGACTGG GAACATTCTG GCCTAGGACA GTGCCTCGCA  
 5601 TTATGTAGGT TCTCAGTAAG CGTGAATGGT GTGTCTGTGT GAGTGGGGGG  
 5651 CCACGAGGCA TGCGCATGTC CAGCAAAGGG CTCACTACCC CTGCCCCCCC  
 5701 AGCCCTACCT ACAAGTGGAA GCGGCAGGTG ACTCAGCGGA ACCCTGTGGG  
 5751 ACAGAAAAAG CGCAAGATGT CCCTGTTGTT TGACCACCTG GAGCCCATGG  
 5801 AGCTGGCGGA GCATCTCACC TACTTGGAGT ATCGCTCCTT CTGCAAGATC  
 5851 CTGGTGCGGC CCGAGGGCTG GGGGGTCAGG GGTCCAATGT GGGCTGGAAG  
 5901 AGAGTTCTAG GAGGGGAGG GTCCCTGGCG TAGGCTGGGT CACAGGGTGC  
 5951 ATCAGGGGTT TCAGTGTAAC CACTGAAGGT CAGCTGGAGG GTGAGGAGTG  
 6001 GCTATCAGTG AGGGGAGAGG CCGGCAAGGT GCTGAGGCCA CTCCTCATGC  
 6051 CCCAGTTTC AGGACTATCA CAGTTTCGTG ACTCATGGCT GCACTGTGGA  
 6101 CAACCCCGTC CTGGAGCGGT TCATCTCCCT CTTCAACAGC GTCTCACAGT  
 6151 GGGTGCAGCT CATGATCCTC AGCAAACCCA CAGCCCCGCA GCGGGCCCTG  
 6201 GTCATCACAC ACTTTGTCCA CGTGGCGGAG GTGCCTGCCC CTCCTCCCG  
 6251 GTGTCTCCA ACCACCCAC ATGCCAGTCA GGCCAACCCT TCCCTTCCCC

FIGURE 3, page 2 of 12

6301 TAACCCACTG CCTTCTCTCT AGATAAGCTG GGCCAAATTC TGGGCCCACT  
6351 CAGTGACTCC CTGCCCTCCT GTCCCCATTT GCCTTCCAGA AGCTGCTACA  
6401 GCTGCAGAAC TTCAACACGC TGATGGCAGT GGTCGGGGGC CTGAGCCACA  
6451 GCTCCATCTC CCGCCTCAAG GAGACCCACA GCCACGTTAG CCCTGAGACC  
6501 ATCAAGGTGC CTGGGACTGG GGAGGGGCGG GTGCTTCCCA GGTCTGTCTT  
6551 CACTGGGTCC TCCCAGCAGC ACTGGGGGCT GGGCACAGCT GTCCTCATTT  
6601 GATAGATATG GAAATGGAGG CTCAGAGGGG TTAAGTGCTT TTCTCAGTTT  
6651 GCACAATGGC AACAGCAGAG TGGGGGCTCA CAGGTCGTCA GGGACCCCAA  
6701 AGCTAGTACT TTTTTTTTTT TTTTAAAGAC AGGGTCTCTC TCTCTGTTGT  
6751 CCAGACTGGA GTTCAGTGGT GCAGTCACAA GCTCACTGCA GCCTTGAATT  
6801 CCTGAGCTCA ATCGATCCTC CCACCTCAGC CTCCTGAGTA GCTGGGACTA  
6851 CAGGTGTACG CCACCATGCC TAATTTTTGT ATTGTTATTA ATTTTTTTTT  
6901 TTTTTTTTTA GAGATGGGGT TTTGCCATGT TGCCAGACT GGTCTTGAAC  
6951 TCCTGGGCTC AAGTGATCCG CCTGCCTTGG CCTCCCAAAG TGCTGAGATT  
7001 ATGGCTTGAG CCATTGTGCC TTGCCACTTG TAGTTTCTTC TTTTCTTTCT  
7051 CCTTCATTTT TTATTATTTT TGAAGTATTT TGAAGTATTG AGTAACATAC  
7101 ATATAGAAAA GTATATAAAA ACATATGAGA CTGGGCGTAG TAGCTCACAC  
7151 CTGTAATCCC AGCACTTTTG GAGGCTGAGG TGGGCAGATC ACGTGACATC  
7201 AGGAGTTTGA GACCAGCCTG GCCAACAAAG TGGAAACCCA TCTCTACTAA  
7251 AATACAAAAA TTAGCCAGGC ATGGTGGCAG GCACCTGGAA TCCAAGCTAC  
7301 TTGGGAGGCT GAGGCAGGAG GAGAATTACT TGAAGTCAGG AGGCGGAGGT  
7351 TGCAGTGAGC CAAGATTGTG CCACTTCACT CCAGCCTGGG CGACAGAGTG  
7401 AGACTCCATC TAAAAAATAA GAAAAGTATA TAAAAACATA TGAATAGTTT  
7451 AAAGAAAAAT TGTAAAGAAA ACACTGTGTA ACTACTGCCC GGGTTGGGAA  
7501 ATAGAACCTT GCCAGGCCCC CAAGCGCCCA GCACTTTAGA GCATAACTCC  
7551 CTCCCACGTA CTTTGGCAAT GATGATCTTG CTTTCTTTTA TAGCTTCACC  
7601 ATGTAGGTAT GCGGTCCAAA ACAATGTGGG GCTTTTTGTT GTCTGTTTGT  
7651 AACTTTCAT GAATGGAATG TTGTTTGTGT TATTTTATGT CTTGCTTTTT  
7701 TCATTCCACA TGGTCTGAG AGTCTTTTCA TTCTGTCTATG TGGAGCAATT  
7751 GTTTTTTCAT TTTCATTGCC ATATAATATT TTATTGTACG TCTACCCCAA  
7801 TTCATTTATT TATTTATTTT TTTGAGATGG AGTCTGTCTC TGTCATCCAG  
7851 GCTGGAGTGC GGTGGCGAGA TCTCATCACT GCAACTTCCG TCTCCTGGGT  
7901 TTACCTGATT CTCGTGCCCT AGCCTCCTGA GTAGCTGGGA TTATGGGCTC  
7951 GTACCACCAC GTCTGGCTAA TTTTTTGTAG AGACAGGCTT TCACCATGTT  
8001 GCCGAGGCTG GTCTTGAACCT CCTGAGCTCA GGCAATCCAC CCGCTTTAGC  
8051 CTCCCAAAGT GCTGGGATTA CAGGTGTGAG CCACTGCCCC CAGCCTACCC  
8101 CAATTTATGT ATTGATTCTA TTGTGGAATG TTGGGGTTTT TCCTTTTCTT  
8151 TTCTTTCTTT CTTTTTCTTT CTTTTTTTCT TTTTTTTGGA GAGGGAGTCT  
8201 TGCTCTGTGC CCAGGCTGGA GTGCAGTGAC GCTAATTTGG CTCACTGCAT  
8251 CACTGCACCC TCTGCCTCCC GGGTTCAAGC GATTCTCCTG CCTCAGCCTC  
8301 CTGAGTAGCT GGGACTACAG GCATGCACCA CCACACCCGG CTAATTTTTG  
8351 TATTTTTTTA GTAGAGATGA GGTTCACACC ATGTGGCCA AGATGGTCTC  
8401 CATCTCTTGA CCTCATGATC CATCTGCCAT GGCTCCCAA AGTGCTGAGA  
8451 TTACAAGTGT GAGCCACCAC GCCAGCTGG TTTTCCAGT TTTTGCTGTT  
8501 TGGACGGGGT GTCTGAGTAT GTTCTTCCAG GTCATTGTCC TGTGCTGCCT  
8551 TGCCTCCCTG AGCCTCTGTT TCTCCTGTTA AATGTTGATG ATTCCCTGCA  
8601 TCCAGGCCTG GTTTAGAGGT GTGGTGCTTT TGGCAGTGAG TATTGCCTTG  
8651 AATTCATGGC AATGAATTCA ATCCCCAGGG GCTGAGAGAG CCAGTCGTGG  
8701 GGGACAGTAA GGGAGGTTTT TACTCTTTCA CCTGTCCCTG ACCCTGACTC  
8751 CTCCTCACCC CCTCCTACAT TTCCAGGGCT GAGGTAGGGA GGATAGTTGT  
8801 GGGGGTATGA CTCTCTGTC CTTTGTCCCC AGCTCTGGGA GGGTCTCACG  
8851 GAACTAGTGA CGGCGACAGG CAACTATGGC AACTACCGGC GTCGGCTGGC  
8901 AGCCTGTGTG GGCTTCCGCT TCCCGATCCT GGGTGTGCAC CTCAAGGACC  
8951 TGGTGGCCCT GCAGCTGGCA CTGCCTGACT GGCTGGACCC AGCCCGGACC  
9001 CGGCTCAACG GGGCCAAGAT GAAGCAGCTC TTTAGCATCC TGGAGGAGCT  
9051 GGCCATGGTG ACCAGCCTGC GGCCACCAGT ACAGGCCAAC CCCGACCTGC  
9101 TGAGCCTGCT CACGGTGAGG AGCAGGGGCG AGGGAGGTGG GGAGCTGGG  
9151 ACCAGGGGTT GACAGTTTCC CCAGGTCTG GCTGTGGGCG TGGCCTGGGG  
9201 CTCTGGGTTT TGGCCAAGAA ACTGAGATCT AGCGTGGGCT CTGGGGTTTG  
9251 GAGTGATGCT TGAGAAGGGG TCCAGGCTCT GGTGGGGCT GTGGACTGAG  
9301 GTCTGATCTC CAGGCTGGTA TGTGGACTGT GGGCAGTTTG AACTGGGCTT  
9351 GGGTCCCGGG TTGAGTTCTG GCAATGGGCT GTGTTCTAGG GCTGGGCCAA  
9401 GCTCTGCATT TGCTGGGCAG GGGTGGTTTC TAAGCATGGC CCTGGGCTCG

FIGURE 3, page 3 of 12

9451 GAGTGAAGTT CTGGGCTTGG CTTTACACTT GGTCTTGGGG TCTAGGGTGG  
9501 GAGTTGGGTT CTGGTTTAGA TCCAGACAAG GTTCTAGACA TTGGGCTGGG  
9551 GCTTAAGTGT TAAGGTTTGG AGTGGATTCT TAGCTGCTTC TGGGCTCTGG  
9601 AGGGGATCAG GGTTGAAATC AGAGCTTCTG GCTGGGTTCC GACCTGGCTT  
9651 CTTCCCTGAC ATCTTGGCAA TATGTTGTGT TCAAGGTTTG GGGCCATGCT  
9701 GTGGTTTGAT CTGTGCGCTG GGATGACATG GGGGTTGCTG TGCTGTGTTC  
9751 TAAGCCAGGC TTTGTCTTGA GTCTAGCTTC TGACCCGAGC TCTGGCTGAG  
9801 CTGTGGCCTC TAGGTCGACC TTTGGCCCTG GGCTCTGTGG CCGTGGGCAG  
9851 GGGCCAGTGG GGGTGATCAG ATCTGTGTGT CCCAGGTGTC TCTGGATCAG  
9901 TATCAGACGG AGGATGAGCT GTACCAGCTG TCCCTGCAGC GGGAGCCGCG  
9951 CTCCAAGTCC TCGGTGAGGG GGTACTCCCT CCTCTCCACT CTGCCCTTCC  
10001 CTCCTGAGAA TCCCAGGATG TGAGGATGGG AAGAGCTCTT AGCAGCCACC  
10051 TCACCATCCT ATCTTGTAGG ACAGAGGCAT CCTGGGGGTA GGGCAGTAGT  
10101 GTTGGGCAGA CTTCCCTCTC CCAGGGATTG CCCTCTCTGT TCCCCGGGGC  
10151 TCTGGGCTCC CCCTGCCTCT GGCCCTAGCT CAGGCCCGAC CATTTCCATA  
10201 GCCAACCAGC CCCACGAGTT GCACCCACC ACCCCGGCCC CCGGTACTGG  
10251 AGGAGTGGAC CTGGGCTGCC AAACCCAGC TGGATCAGGC CCTCGTGGTG  
10301 GAGCACATCG AGAAGATGGT GGAGGTGAGC TCCTGCGGAG CCTGAGCAGT  
10351 GTGTGGGGAG AGGCCAGTTT GCCGGAGCAC TGCCCTGGAA GCCAGCACGA  
10401 GTGTCTCTGT CAAGACCCAG CACTCAGCCC CTAGGAGTCA CAGGGCCTGG  
10451 CAGGCCAGCT GCACGGGGCT GAAGTGCCCC TGGGTAGGGT GGGGGTGGAG  
10501 GTATGGAACG GGGGTGGTGT CAGAGACCTC TCTGAGACAC ACCTCATCAA  
10551 ATGGACTGGG AACGTGGGAA GGGACAGGAC CTGATGTCCC CTTTACTCTC  
10601 CCCTCTTCTG GCTCTGCGTG TCCCTCTGCG TGCCCCAGTC TGTGTTCCGG  
10651 AACTTTGACG TCGATGGGGA TGGCCACATC TCACAGGAAG AATTCCAGAT  
10701 CATCCGTGGG AACTTCCCTT ACCTCAGCGC CTTTGGGGAC CTCGACCAGA  
10751 ACCAGTGAGG AGGGCTGGGG ACCTGGGGGA GAGGGAAGGC AACTCAGCCC  
10801 ACTTCTGCTT GGGCTTCAGT TTCTTGTGTG CAAGATGAGG TCACTGAGCC  
10851 AGATGATCTT GGCCTGGGAA GCTGCCAGTG TGGGAAAGGG CACTTGCTTT  
10901 TGTGGGGAGG AGAGGCTGCC AGCTGTGGAG GCGCAGTGGT ATCTCACAAA  
10951 TTCAGACAGA TGGGGGGCTC CACCTGAGTC TTGCAAAGAC TGTGACCTGG  
11001 GGACTGTGGC TACAAAAGTG CTGTTTTATT TGTGGAGCTC ACAGCTGTCA  
11051 AGAATAGTGG GCAACTTGAG CTCCTGGATA GTCTGTTCTA ATGAATAGAT  
11101 AAGAAAGGTT TGTAAATTAG AGTACCCAGT TGTTTATCAA CAGTTCATAT  
11151 GCTGACAATT TGGA AAAACA GCTGGTTCTC TGAAGTAGGT TAAACATGCC  
11201 CCCTGAAGCC AGATTTCATGC CCTATTTTTC CTGAGCAGAA AAAACTCCAT  
11251 TCAAAATTTA AAGTCCATCT CAGGTCGATT TATTTTTTAA TGTTACCTGT  
11301 ATTTCAAAA TCTGTTGTTT TTTATTTCCA CATTACAAA ATCCACGGTA  
11351 AAATAAAATC TAGGTGGTAA AATAAATTTA TAGTGAACAA AATGTTTTAA  
11401 GTAAGAAGTG AGAGGCCAGG TGCGGTGCCT CACGCCTGTA ATCCTAGCAC  
11451 TTTGGGAGAC TGAGTTGGCA GGATCAATTC AGGCCAGGAG TTTGAGCCCA  
11501 GCCTGGGCAA CAGAGTAAGA CCCTGTCTCT ACAAAAATTA TTATTATTAT  
11551 TTTTGAGACA GAGTCTCACT CTGTTGCCCA GGCTGGAGTG CAGTGGTACA  
11601 ATCTCGGCTC GCTGCAACCT CCACTTCCTG GGTTCAGTG ATTCTCCTGC  
11651 TTCAGCTTCC TCGAGTACTG GGATTACAGG CATGCATCAC CGTGCCTGGC  
11701 TAATTTTGTG ATTTTGTAGCA GAGATGGGGT TTTACCATGT TGGCCAGGCT  
11751 GGTCTCAAAC TCTTGACCTC AAGTGATCTA CCTGCCTTGG CCCCCAAAG  
11801 TGCTAGGATT ACAGGCATGA GCTACTGCTC CTAGCCTAAA AAAATTTTTT  
11851 TTGGGCATGG GTGGCACGTG CCTGTAGTCC CAGCTACTCA GGAGGCTGAG  
11901 GCAGGAGGAA CCCTTGAGCC CAGGAGGTTG AGACTGCAGT GAGCTGTCTC  
11951 CACACCACTG CACTTCAGCC TGGGTGACTG CGCGAGATCA CCCCCATCAA  
12001 AAAAAAAAAA AAAAGAAAAA AAAAGGAAGA AATGAAAGTC CCCTCTTTCC  
12051 TTTTCCACTG GTAGAAGTTG CCATGATTAA GCACTGTAA CAATATTAAG  
12101 CTTGGCAGTA TGTGGATTCT TCCAGTCTTC TTTTCCAGG CAGGTGCACA  
12151 TTGATAGAGA TTTTGTGTTG TTGGTGTCTG TTTTCATGGAC AAACAGGATT  
12201 AGAGCATAAA TCTAGTCTCT CTTGTGGCTT TTATCATAGC TGCTTTATTT  
12251 CTTCTCCAG ATTTTAGGCA GAGGTAGTTG AGTTCCATGT TTTCTCCCTG  
12301 GGTTGGTGGG TGGATTTTTA TCTAGACCAC CTTTTCAGTG AGAATGACCC  
12351 TTTGAGACGA TGGAGGCCTC AGCTTCATGC AGCGGGCTCA GCCTTAACCC  
12401 TCCACCTCCT GCAGGCCCA AGCTGTGTGT GTGTGTGTGT GTGTGTGTGT  
12451 GTGTGTGTGT GTGTGTGTGT GTTGGTAAGG GGAAAGCCCC TGTTGGGTA  
12501 TCAAAAACCT AGCACCTGGT TCGGCAGGAG GGAGACCAGC ACCGGCTCCC  
12551 CAGGACCAGG CCGAGCTCAC CACTTCATTG TAAAGCTCCC TCTTTGTTTC

FIGURE 3, page 4 of 12

12601 TGGAACCTGG GTGTTTCCAT TTCTTTCTTA CAAAATTATC TATGCATTTA  
12651 CAGCAATTGT TGATATATCT TTAGGCAGCA TCTAGGTA CTGTAGTGGGT  
12701 TCTCTTTTTT CTTTTTCTT TTTTTTAATC ACCCTCTCTT TTTTTTGAGA  
12751 CAGAGTCTCA CTCTGTCGCT CAGGCTGGAG TGCAATAGCG CGATCTTGGC  
12801 TCACTGCAAC CTCTGCCTCC CAGGTTCAAG TAATTCTCAT GCCTCAGCCT  
12851 CCCAAGTAGC TGAGATTACA GGCAGTGGCC ACCAGACCCG GCTAATTTTT  
12901 TTTTCTTTTT CTTTTTTTTG AGACGGAGTT TCGCTCTTTG TTGCCCAGGC  
12951 TGGAGTACAG TGGTGTGATC TCGGCTCACT GCAACCTCCG CCTCCCGGGT  
13001 TCAAGTGATT CTCTGTCTC AGCCTCCCGA GTAGCTGGGA TTACAGGCGC  
13051 GCGCCACCAT GCCTGGCTAA TTTTGTATTT TTTTTTTTTT GAGACAGAGT  
13101 CTCACCTGTG CACCCAGACT GGAGTGCCTG GCGCGATCT CGGCTCACTG  
13151 CAAGTCTGCT TCCCGGGT CATGCCATTC TCCTGCCTCA GCCTCCGGAG  
13201 TAGCTGGGAC TACAAGCACC CACCACCGTG CCCGGCTAAT TTTTTGTATT  
13251 TTTAGTAGAG ACGGGTTTC ACCGTGGTCT CGACCTCCAG ACCTCGTGAT  
13301 CCACTAGCCT CAGCCTCCCA AAGTGTGGG ATTACAGGCG TGAGCCACCT  
13351 CACCCAGCCT AATTTGTAT TTTAGTAGA GATGGGGTTT CACCATGTTG  
13401 CGCAGGCTGG TATTGAACCT CTGACCTCAG GTGATCCGCC CGCCTCGGCC  
13451 TCCCGAAGTT CTGGGATTAT AGGCGTGAGC CACCGCACCT GGCCTAATTT  
13501 TTGTATTTTT AGTAGAGATG GAGTTTTACC TTGTTGGCCA GGCTGGTCTT  
13551 GAACTCCTGA CCTCACCTCA GGTGATCTGC CCACCTCGGC CTCCCAAAGT  
13601 GCTGGGATTA CAGGCATGAG CCACTGTGCA CCCGGCCTAA AAATCACCAT  
13651 CTTGACAGAA CTTACGCCT TGCTTTTTGT TTTTTTTCAT CTTTGTGCTT  
13701 GTTTTCCACT TAACCCTTGA TCACAGACAT CTTTCCATGT GGATTCATGT  
13751 AGAACTACCT CATTCTTAG AACAGCTGCA GAGTATTCCA CTGTGCGGTT  
13801 AGTCCATCAT TTCCCTAACC ATCTCCTGC TGATGGACAG TTAGACTGTT  
13851 CCAGTTTTTC AGTATGATTC TATGCCAGGC TGCCATGAAC GTCCTTTTAC  
13901 TGATCCACTC AGGCCAGTAT TTCTGTAGGA GAAATTCTTA GAAGTGGGAT  
13951 AATTGGATCA AAAGATATGC ACATTCTAAA TTAGGAGAGA GACTGCCAAA  
14001 CTGACCTCAG ACAAGGTTGT ACCAGTTTGC ACCCCCATCA GCAGCGTACA  
14051 AGTGCCTGCT TCCCAACTTC CTCGCCAACA GGGATGCTAT AAAAAGCTTC  
14101 ACAATTTTGC CAGTCTCATT GGCAAATGGT ATCTTGGTTA AATTTGCATT  
14151 TCTTTAATAC TAAGTGGGGG TAGGGTATCT TTTTATATGT TTATTGGCCA  
14201 TTTATTTCTT CTGTCAATTG CCTGTTCTGA TTCCTTGTCC ATTATTCTAC  
14251 TGGGTTTGTG GGTCTTTTTT TCATTGATTT TTAGAATCTC TGTAAATGGA  
14301 TATTAACCTT TTGCTGTTGA ATGTGTTTGC AAATATTTTC TCCCTGTCTG  
14351 TCATTTATGT GTCTTTTTTCC ATATAAATTT AAAAAATTTT GGTGGGCTCA  
14401 ATAGGTCAGT CTTTCCCTTC CGGGCTTCTG GGATTTGTGT TCGGGGTAGA  
14451 AAGGCCCTCA GCCCCTCAAG ATTATAAAAT TATAAAACCT TTTCTTTTTT  
14501 TTTTTTTTTT CTGAGACAGG GTGTCTTGCC ATGTCACCCA GGCTGGAGTG  
14551 CAGTGGCATG ATCTTGGCTC GCTGCAACCT CCACCTCCCA GGTTCAAGTG  
14601 ATTCTCGTGC CTTAGCCTCC CGAGTAGCTG GGATTATAGG TGCTGCCAC  
14651 TATGCCTGGC TAATTTTTTG TATTTTTAGT AGAGACGGGG CTTTGCCATG  
14701 TTGGCCAGGC TGGTCTTGAA CTCCTGACCT CGTGATCCAC CCGCCTTGGC  
14751 CTCCCAAAGT GCTGGGACTA CAGGCGTAAG CCACTGTGCT CGGCCCTATA  
14801 TTTTTTTCAG ATAGCCAGTT ATCCTAATGC TCCCTTGATT TGATGGACCA  
14851 CCTGGATCAC ACATTATGAG CCCCCTCATA AGCAGGTGGG AGTCTCAAGC  
14901 GAGGGCCAGT CCCGATGGGA ATAGCACTTG GTGGCTGAGG ACCCTCCTAT  
14951 CTGTGCAGAC ACTGTTGTAA AACTTCACAT GCATCATCTA ATTTAGTCCT  
15001 CACCAAAATC CTATGAAATG TAGGAATGAT CATTACACCC ATTTATAGAT  
15051 AAGGAAACGG AGGGACAGGG AGATTACTCC GCTACAGGTC AAGAGGCAGG  
15101 GAAGTAGAGC TGGGATTTGA ACTGAGGTCT GTGTCTAGAA CACGTGCTCA  
15151 TTCTTTCCCT AAAATGTATT CATAGGTGAA AAAGGGCTTC TGCGGAAAGC  
15201 CCTGGGTTAT GTGGGAAACC CTGGATTTAC AGCTGTCTTT CCAGCAGGAT  
15251 GATGCAGGAG AGAGAGGGAT GCGATTTCTC CCAATCTCTC CTGGTCCCAG  
15301 AACTCATTAG AGAGTTCTCC CTGCTGAGGG CTCCCGACTG GTGTTGCACA  
15351 CAGTACACTT CGGGAGCCCG AGGCTGATGG TTCCATGGAA AGTACACAGT  
15401 CATTTTAGTT TGCACACCAA GTGTGAAGTG GGCAGGACAG GCCACTGTTC  
15451 TGAGAAGGAA CCCAGGGAAA GGGACTGGCC CAAGACCACA CACTGGTTAG  
15501 CGGCACTTCC CACATCTGCC TGACCCCTAG TCCAGTGCCG CTTTTTCTTT  
15551 ACTCTGCAAC AGGAGTCCAA AATCAGGAGT TCCATGAGGA CACTGGGAAC  
15601 AGTGGGATGG GTTAGGCCAG CGGTGGATGG TTCTGGGGAG GGCCCGAGCT  
15651 GAAGCGCCCC CGCAACTCCC CACAGGGATG GCTGCATCAG CAGGGAGGAG  
15701 ATGGTTTCTT ATTTCTGCG CTCCAGCTCT GTGTTGGGGG GGCGCATGGG

FIGURE 3, page 5 of 12

15751 CTTTCGTACAC AACTTCCAGG AGAGCAACTC CTTGCGCCCC GTCGCCTGCC  
15801 GCCACTGCAA AGCCCTGGTG AGAGTCCCTT TCCCGGCTCA CGGCCCAAGC  
15851 CACGCCCTC CAGCCCCGGC CCCGCCCTCC CTTCTGGCCC CGCCTCTGCC  
15901 AGAGCCCTTC TCAAGCCAGG AAAACCTGGT AATTCTATTT GCCTCTCCTC  
15951 CTGTGGTTCT GCCCCGGGCC CTGAGGCGGG CTCTAAAGCC CTAGTCTCAC  
16001 CCTCAAGAAG GAAGAAGTAG AGTCATCACC TCTAAATCCC TCCTCCCAAC  
16051 ACGGCCCTC CTCTATTGCA GATCCTGGGC ATCTACAAGC AGGGCCTCAA  
16101 ATGCCGAGGT GAGATGGAAT GACTGGAAGG GCTGCTGGGC AGTGTTTTTT  
16151 TTGTTTGTGT GTTTGTGTGG GAGAGTTACT ATTTTGGTGG GGCAATGGCC  
16201 AAGGAGTGAA GTACCTTAAA ATCAGAGGCG CATGGCCGGG CATGGTGGCT  
16251 CAAGCCTGTA ATCCCAGCAC TTTGGGAGGC CGAGGCGCGC AGATCACCTG  
16301 AGGTCAGGAG TTCAAGACCA GCCTGACCAA CATAGCGCAA CCCCCTCTCT  
16351 ACTAAAAATA CAAAAAGTAG CTGGGCGTGG TGGCACCCAC CTGTAATCCC  
16401 AGCTACTTGG GAGGCTGAGG CATGAGAATC GCTTGAACCT GGGAGGCGGG  
16451 GTTTGCAGTG AGCCGAGATC ACGCCACTGC ACTCCAGCCT GGGCAACAGA  
16501 GAGGGCTCTG TCTCAAAAAA AAAAAACAAC AAAAAAACCC CCAAAACCAA  
16551 AACCACACAA AATCAGAGGC TCAAGATGAC TGATGTGAAG GGAGTGGCGT  
16601 TTAAGAGGCC ATTTATTTTG ATGACGCAGC TGCCAGGAA CAGAGAACAT  
16651 GGGAGAAGGC ATAGACTGAC AATTAGGAGG AGGAGAACAC TTTGGAAGGA  
16701 GACTCTTATT TTGGTGGGCG AGCTGCTCAG GAACAAAGGT TCCTGGTAGG  
16751 GGGGCGCAAG CCTGCGGGAT GGGATGGAGG GTATTCTGAC CAATGTCCCT  
16801 GGCTGGCTCT CCATTTGCTC TCCCCAGCC TGTGGAGTGA ACTGCCACAA  
16851 GCAGTGCAAG GATCGCCTGT CAGTTGAGTG TCGCGCAGG GCCCAGAGTG  
16901 TGAGCCTGGA GGGGTCTGCA CCCTCACCTT CACCCATGCA CAGCCACCAT  
16951 CACCGCGCCT TCAGCTTCTC TCTGCCCCGC CCTGGCAGGC GAGGCTCCAG  
17001 GCCTCCAGGT AAGAGGGAGT CATTCTGTAC TGGCCTGTGG AGGGAAGGAT  
17051 GCAGGGCTAC TGGGGCAAG AACGCAGGAT GGAAGCCATT CCAAAGTGCA  
17101 TAATTCTCTT TTTGTGGTGG GATAATAAAG AAGGGACAGG CCGGGCGCGG  
17151 TGGCTCACGC CTGTAATCCC AGCACTTTGG GAGGCCGAGG CGGGCGGATC  
17201 ACGAGGTCAG GAGATCGAGA CCATCCTGGC TAACACGGTG AAACCCCATC  
17251 TTTACTAAAA ATACAAAAAA AAAAAATTAG CCAGGCGTGG TGGCGGACGC  
17301 CTGTAGTCCC AGCTACTTGG GAGGCTGAGG CAGGAGAATG GCATGAACCC  
17351 GGGAGGCGGG GCTTGCAGTG AGCCGAGATC GCGCCACTGC ACTCCAGCCT  
17401 GGGCGATAGA GCAAGACTCC GTCTCAGAAA AAAAAAAAT AAAAAATAAA  
17451 GAAGGGACAG GTAAGGGTGC CAGAAAGTGG CCAGGAAGCC CTGGACCTTC  
17501 TGAGGCTGAG GAGAGAGACC CTAATTTATA AAGAGGTATA AAAGTGAAAG  
17551 AGGCTTCAAG ATTCCAGTTA CAGTCTTATT TTGTTGGAGG GGTTAACAAA  
17601 GGATTGGAGA AGGTGTTATA TGAGCCATTG GCTTGCCTTT CCCTTTCTGG  
17651 CTGCTCTGGA GGCTCTTCTG GGGAAAGTCC CTTGCCCTGA TAATGTCTTG  
17701 GCAGCTCTCT TGGGCTATTG GATGGTTTGA GGTCAGTTTG CTGAATGACA  
17751 ACTGGCCAAA TGATTATTTT GCTGAGAACG GTCCGAACAA CTATGTTAAA  
17801 CTGGGGTCTA AGGTAGTTGA TCACAACTGT TTGGGTGGC ATAAGTCCCTC  
17851 AAAAAACAGA GGCAGGCACA GGGCATACAT CCTCAAAAAT AGAAAAGATA  
17901 AATCCATTTG CATTGAGCCT TCCAGAAGTG CTGGGGTCTA AAATGTGAAA  
17951 TACACACAAA ATTGACATTT AAGCAAACTG CGCTGACAAA TCTGTGGCTG  
18001 AAAAAGCTGT GGCAAAAACAA AAACATAGAA AAAGAGCCTC AAAAATTGGG  
18051 CTGAGGCCAG GCATGGTGGC TCACGCCTGT AATCCTAGCA CTTTGGGAAG  
18101 CCAAGGTGGG TGGATCACCC GAGGTCAGGA GTTGGAGACC AGACTGGCCA  
18151 ACGTGGCAAA ACATCATCTC TACAATACAA AAATACAAA ATTAGCTGGG  
18201 CGTGGTGGCA GGCCTGTGTA ATCCCAGCTA CTTGGGAGGC TGAGGCACGA  
18251 GAATCGCTTG AACCTGGGAG GTGGAGGTTG CAGAGAGCCG AGATTGCGCC  
18301 ATTGCACTCC AGCCTGGGCG ACAGAGAGAG ACTCTGTCTC AAAAAAAAAA  
18351 AAAAAAAAAA AAATTGGGCT GTGAGGTCAT GCAGGGAATT GATTTTGGT  
18401 GGGTGGGTCT GCTTCTGGGA TGATGTGGAT GCCTCCCGTG GAGAGGGGAA  
18451 GGGTTGATGA AGTCCAGGG ACCTGGAAGT GTGTTCTGCA GCAATCCCCC  
18501 TCCCAGCAGA GATCCGTGAG GAGGAGGTAC AGACGGTGA GGATGGGGTG  
18551 TTTGACATCC ACTTGTAAATA GATGGTGAGT CCTCCCACAG CTGGCACCAG  
18601 AGTCCCCAC TGAGGGCTGG GGGGAGCTG GGGAGTATCA GGGAAATGGG  
18651 TGCTTTATCC AAATGGCTCC AAGCCAGGTG GGCTACTACC TTGTTGTTAG  
18701 GGGGGTGTCT TCCTCACAACT CTGTTTTTCT CTTCCAGCT GTGGTTGGAT  
18751 CAAGGACTCA TTCCTGCCTT GGAGAAAATA CTTCAACCAG AGCAGGGAGC  
18801 CTGGGGGTGT CGGGGCAGGA GGCTGGGGAT GGGGGTGGGA TATGAGGGTG  
18851 GCATGCAGCT GAGGCGAGG CCAGGGCTGG TGTCCCTAAG GTTGTACAGA

FIGURE 3, page 6 of 12

18901 CTCTTGTGAA TATTTGTATT TTCCAGATGG AATAAAAAGG CCCGTGTAAT  
 18951 TAACCTTCAC CATCAGCGCC TAGAATCCCG GGGGGTAGGG GGATGGTATA  
 19001 CTTTACAGGA TGACAATCTT GGGAGCTAGA ACTTTGTAGC CAGAGAACT  
 19051 TGGGAGGTCT GGAATCTCAT GTGTCTGGAG TCTTGGGGAA GAGAATCTTA  
 19101 GAAGCAGAAA ACCTTGGAAC ATAAGAATCT TGGGGAGGGT CTAGGATCTT  
 19151 GAGGAGACCA GATCCTTGGA CATCTAAAAC TTGAAACTAG TAGGTCTGCA  
 19201 CCCGAGAATT GCAGGGCCAG TCATGCATAC CCAAAGCCTT CAGCCCATGG  
 19251 CCGAAATTCC CTTGCTGGAC AGGGGGCCTT TCAGCCCCTG CTTGGACGCT  
 19301 TCCAGTAACA GGGCCCTCAC TGCAGGAATC GTGGGAGGGA GAGGGGCAGC  
 19351 ACAGAGTTGC TGGCTGTCGG GGAAGGGAGG GAGGGCCCTG GGCAGTCCGA  
 19401 GGGCCCTGCT GGGCTTGTGC CTCAGGGTGG GGGCTGCACT CCTCCGCCTT  
 19451 GCAGCCTCCT GGCCTGGTGC TGCTGCCAGC CGGAAGGACA GTGACTTCCA  
 19501 GAGGAAATGC ATATTGATCC TGCTTTCAGC CTCGGTGGT GGCTTCTCCC  
 19551 AACCCAGCTC TTCCCTCCTG AGCCTGCAGC ACGGAGGTTT TGGGGGTCAC  
 19601 TGCTACCTAA AGAAGGCTAA GGCCACTTCT GAGGCTGGTC TGGGAGTTTA  
 19651 CTAAAGGTTC TGAAGCTGGG CCGGGCTGCC CCTGGGATCA GGAGACTCCA  
 19701 GACAGCAGTC CTGACAATGG GAACTACCTC CTCAGTCCCC CAAACTGGGA  
 19751 GGTGTCCAC AGCAGCTGTA GGATTGTCCT AGGGGTGGAG ACCTGAGCAC  
 19801 CTTCCACTCC AAAGCAGAT ATCTGTGGC CTGGCAGTGG CCTCAGTTCC  
 19851 CCCATGAGTG CCCCCTGCC CCACCCAGG GTTTCCCCAC ATCACATCCA  
 19901 TCCCTGCTTT GAGACCCAC TCCCCCTGGC CTGTTCTTTA TTTTGGGTCA  
 19951 CTCCCTTCTC TTTCTGGTC ATATCTCTCC TGCAGGCCTA CCCTGTGTGTG  
 20001 GGCCCCCAG CCCTGTCTCT GCATCGGGTG CCCCCCTGCC CCTCCTTCTG  
 20051 TCCTCAGCCC CCTCCGCCCT TCCCCCTCTT GAGGCTGTAA TATCCGTTTC  
 20101 ACGATTGGG GGCTGAGTTG CTATAACAAC AGACGGCGAT TGTGTTGTGA  
 20151 AGAGCAGCTC GCTCCTGTGC CGCCTGCCTC CTGTGCTGCC TCCATCCCTG  
 20201 CAGCCCAGTC GGTTCCCTCT GGCTCCTCTC GTCACTACCC TCCAGTTCCA  
 20251 GTCTGGCCTC TTCCTGGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTAT  
 20301 GCATGCATGC ATATGTGTGT CCAGGTCTGC CTGCCCAGGA TGTGACAAGT  
 20351 AGCGGTCTTC ATGGTTGCAT GTGTCTGAAT TTGGTGTCTG AGCTTCACAT  
 20401 TGTATGCGCC TGTGTGCATG TGTGTGCATG GACATGCATG CTGTATCTGC  
 20451 TGTGTTTCCC CTCCCCCATG TGTCCCCACT GGCCTTTGCA CATGGGAGAA  
 20501 GGGCATGTGC TCAGCATATC ACTCAACTGT CCACATTGGG TGGGTACCTG  
 20551 TGTGTGGTGT TGTGTGTGG GGGGTGTGTC TTGAAGTGGC AGGTCCCAAA  
 20601 TGCTTAGGCA ATCTGAACCT TGGACCTTGC AGAGAGGAGA GATGTCCCTG  
 20651 TAGGTGGGAG GGACAGGGAG ATGCAGCAGC TGCCCGGTGA CCTTTTCTGC  
 20701 CCTTGATGGG CAAAGCTGGG GGTAGGGAAA GGAGACAAGT GCTCATACTT  
 20751 ACCTCCCTCC CTGCCCAGGC TCCTCTGTAA GGGTCTGAGT CTGTCTCTGT  
 20801 GAGCCATTGC ATCTGTCTGT CTATGCCCTG ATGCCTGGAT GGACAAGGGG  
 20851 TGTGTGTGTG TGTGTGTGTG TGTGTGTGTG AGTGTGAGGC TGCAGGAAGA  
 20901 GGAACAGTGG GGGATGGGCA GGAAAGTGGG CTGTGGGGTC AGGGAGGCGA  
 20951 T (SEQ ID NO:3)

# **FEATURES:**

Start: 3000  
 Exon: 3000-3072  
 Intron: 3073-3753  
 Exon: 3754-3855  
 Intron: 3856-4363  
 Exon: 4364-4427  
 Intron: 4428-4786  
 Exon: 4787-4918  
 Intron: 4919-5702  
 Exon: 5703-5853  
 Intron: 5854-6056  
 Exon: 6057-6230  
 Intron: 6231-6389  
 Exon: 6390-6506  
 Intron: 6507-8832  
 Exon: 8833-9114  
 Intron: 9115-9885  
 Exon: 9886-9963

FIGURE 3, page 7 of 12



Intron: 9964-10201  
Exon: 10202-10324  
Intron: 10325-10638  
Exon: 10639-10754  
Intron: 10755-15675  
Exon: 15676-15817  
Intron: 15818-16071  
Exon: 16072-16108  
Intron: 16109-16828  
Exon: 16829-17008  
Intron: 17009-18491  
Exon: 18492-18565  
Stop: 18566

**CHROMOSOME MAP POSITION:**  
Chromosome 11

**ALLELIC VARIANTS (SNPs):**

DNA			
Position	Major	Minor	Domain
5539	C	G	Intron
5658	T	G	Intron
5861	C	T	Intron
6023	A	G	Intron
6799	C	T	Intron
9579	C	A	Intron
9842	T	C	Intron
10159	T	C	Intron
12025	A	- G	Intron
14723	T	C	Intron
14996	G	A	Intron
16153	T	G	Intron
16181	G	A	Intron
16756	A	G	Intron
18059	A	G	Intron
18364	A	- T	Intron
18861	G	A	Beyond ORF(3')
20443	G	A	Beyond ORF(3')
20881	A	T	Beyond ORF(3')

Context:

DNA

Position

5539 AGACTCAGTTCAGGCATGAAGTCTCCGTGGGCTCTGAGGGTTCGGGGCTCTTCCGGGGTA  
GAATTTGTCGTTCCCACCTCTGTTTTCCATGGCACTTTGTACAGACTCCTGTACAAAGAC  
CTCTGTACATGTGTACGCTGTTTTGTGATCATGTGTTTCTGTGTCTGTCTCCCTCAGTA  
GACTGTGAGCTCCTCGAGGGCAGGAACCGTGTCTTACTCATCTCTGTATTCCCAGCGCCT  
AGCACAGTGCCTGGCACAGAGTACGTTGTTCAATAATGTGTGTTGAGTGCATGACGGGGT  
[C,G]  
GGGGGAGATGAGGAGGAGTTGCTGGGACTGGGAACATTTCGTGCCTAGGACAGTGCCTCGC  
ATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTCTGTGTGAGTGGGGGGCCACGAGGC  
ATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCAGCCCTACCTACAAGTGGA  
AGCGGCAGGTGACTCAGCGGAACCCGTGTGGGACAGAAAAAGCGCAAGATGTCCCTGTTGT  
TTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCT

5658

CCTCTGTACATGTGTACGCTGTTTTGTGATCATGTGTTTCTGTGTCTGTCTCCCTCAGT  
AGACTGTGAGTCTCCTCGAGGGCAGGAACCGTGTCTTACTCATCTCTGTATTCCCAGCGCC  
TAGCACAGTGCCTGGCACAGAGTACGTTGTTCAATAATGTGTGTTGAGTGCATGACGGGG  
TGGGGGGAGATGAGGAGGAGTTGCTGGGACTGGGAACATTTCGTGCCTAGGACAGTGCCTC  
GCATTATGTAGGTTCTCAGTAAGCGTGAATGGTGTGTCTGTGTGAGTGGGGGGCCACGAG

FIGURE 3, page 8 of 12

[T, G]  
CATGCGCATGTCCAGCAAAGGGCTCACTACCCCTGCCCCCAGCCCTACCTACAAGTGG  
AAGCGGCAGGTGACTCAGCGGAACCTGTGGGACAGAAAAGCGCAAGATGTCCCTGTTG  
TTTGACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCC  
TTCTGCAAGATCCTGGTGCGGCCCCAGGGCTGGGGGTGAGGGTCCAATGTGGGCTGGA  
AGAGAGTTCTAGGAGGGGCAGGGTCCCTGGCGTAGGCTGGGTACAGGGTGCATCAGGGG

5861 CTGGGACTGGGAACATTCTGTGCCTAGGACAGTGCCTCGCATATGTAGGTTCTCAGTAAG  
CGTGAATGGTGTGTCTGTGTGAGTGGGGGGCCACGAGGCATGCGCATGTCCAGCAAAGGG  
CTCACTACCCCTGCCCCCAGCCCTACCTACAAGTGAAGCGGCAGGTGACTCAGCGGA  
ACCCTGTGGGACAGAAAAGCGCAAGATGTCCCTGTTGTTTGACCACCTGGAGCCCATGG  
AGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCTTCTGCAAGATCCTGGTGCGGC  
[C, T]  
CGAGGGCTGGGGGTGAGGGTCCAATGTGGGCTGGAAGAGAGTTCTAGGAGGGGCAGGG  
TCCCTGGCGTAGGCTGGGTACAGGGTGCATCAGGGGTTTCAGTGTAACCACTGAAGGTC  
AGCTGGAGGGTGAGGAGTGGCTATCAGTGAGGGGAGAGCCGGCAAGGTGCTGAGGCCAC  
TCCTCATGCCCCCAGTTTCAGGACTATCACAGTTTCGTGACTCATGGCTGCACTGTGGAC  
AACCCCGTCTGGAGCGGTTTCATCTCCCTCTTCAACAGCGTCTCACAGTGGGTGCAGCTC

6023 GGCAGGTGACTCAGCGGAACCCCTGTGGGACAGAAAAGCGCAAGATGTCCCTGTTGTTTG  
ACCACCTGGAGCCCATGGAGCTGGCGGAGCATCTCACCTACTTGGAGTATCGCTCCTTCT  
GCAAGATCCTGGTGCGGCCCCGAGGGCTGGGGGTGAGGGTCCAATGTGGGCTGGAAGAG  
AGTTCTAGGAGGGGCAGGGTCCCTGGCGTAGGCTGGGTACAGGGTGCATCAGGGGTTTC  
AGTGTAACCACTGAAGGTCAGCTGGAGGGTGAGGAGTGGCTATCAGTGAGGGGAGAGGCC  
[A, G]  
GCAAGGTGCTGAGGCCACTCCTCATGCCCCCAGTTTCAGGACTATCACAGTTTCGTGACT  
CATGGCTGCACTGTGGACAACCCCGTCTGGAGCGGTTTCATCTCCCTCTTCAACAGCGTC  
TCACAGTGGGTGCAGCTCATGATCCTCAGCAAACCCACAGCCCCGAGCGGGCCCTGGTC  
ATCACACACTTTGTCCACGTGGCGGAGGTGCCTGCCCCCTCCCTCCCGGTGTCTCCCAACC  
ACCCACATGCCAGTCAGGCCAACCCCTCCCTTCCCCTAACCCACTGCCTTCTCTCTAGA

6799 CCATCAAGGTGCCTGGGACTGGGGAGGGGCCGCTGCTTCCAGGTCTGTCTTCACTGGGT  
CCTCCCAGCAGCACTGGGGCTGGGCACAGCTGTCTCATTTGATAGATATGGAATGGA  
GGCTCAGAGGGGTTAAGTGCTTTTCTCAGTTTGACAATGGCAACAGCAGAGTGGGGCT  
CACAGGTCTGTCAGGGACCCCAAAGCTAGTACTTTTTTTTTTTTTTTAAGACAGGGTCTC  
TCTCTCTGTGTCCAGACTGGAGTTCAGTGGTGCAGTCACAAGCTCACTGCAGCCTTGAA  
[C, T]  
TCCTGAGCTCAATCGATCCTCCCACCTCAGCCTCCTGAGTAGCTGGGACTACAGGTGTAC  
GCCACCATGCCTAATTTTTGTATTGTATTAAATTTTTTTTTTTTTTTTAGAGATGGGG  
TTTTGCCATGTTGCCAGACTGGTCTTGAACCTCTGGGCTCAAGTGATCCGCCTGCCTTG  
GCCTCCCAAAGTGCTGAGATTATGGCTTGAGCCATTGTGCCTTGCCACTGTAGTTTCTT  
CTTTTCTTTCTCCTTCAATTTTTTATTATTTTGAAGTATTTTGAAGTATTGAGTAACATA

9579 CTGGTTGGGGCTGTGGACTGAGGTCTGATCTCCAGGCTGGTATGTGGACTGTGGGCAGTT  
TGAAGTGGGCTGGGTCCCGGGTTGAGTTCTGGCAATGGGCTGTGTTCTAGGGCTGGGCC  
AAGCTCTGCATTCTGTGGGCAGGGGTGGTTTCTAAGCATGGCCCTGGGCTCGGAGTGAAG  
TTCTGGGCTTGGCTTTACACTTGGTCTTGGGGTCTAGGGTGGGAGTGGGTTCTGGTTTA  
GATCCAGACAAGGTTCTAGACATTGGGCTGGGGCTTAAGTGTTAAGGTTTGGAGTGGATT  
[C, A]  
TTAGCTGCTTCTGGGCTCTGGAGGGGATCAGGGTTGAAATCAGAGCTTCTGGCTGGGTTT  
CGACCTGGCTTCTTCCCTGACATCTTGGCAATATGTTGTGTTCAAGGTTTGGGGCCATGC  
TGTGGTTTGATCTGTGCGCTGGGATGACATGGGGGTGCTGTGCTGTGTTCTAAGCCAGG  
CTTTGTCTGAGTCTAGCTTCTGACCCGAGCTCTGGCTGAGCTGTGGCTCTAGGTGAC  
CTTTGGCCCTGGGCTCTGTGGCCGTGGGCAGGGGCCAGTGGGGGTGATCAGATCTGTGTG

9842 TGGGCTGGGGCTTAAGTGTTAAGGTTTGGAGTGGATTCTTAGCTGCTTCTGGGCTCTGGA  
GGGGATCAGGGTTGAAATCAGAGCTTCTGGCTGGGTTCGACCTGGGCTTCTTCCCTGACA  
TCTTGGCAATATGTTGTGTTCAAGGTTTGGGGCCATGCTGTGGTTTGATCTGTGCGCTGG  
GATGACATGGGGGTGCTGTGCTGTGTTCTAAGCCAGGCTTGTCTCTGAGTCTAGCTTCT  
GACCCGAGCTCTGGCTGAGCTGTGGCCTCTAGGTGACCTTTGGCCCTGGGCTCTGTGGC  
[T, C]  
GTGGGCAGGGGCCAGTGGGGGTGATCAGATCTGTGTGTCCAGGTGTCTCTGGATCAGTA  
TCAGACGGAGGATGAGCTGTACCAGCTGTCCCTGCAGCGGAGCCGCGCTCCAAGTCCTC

FIGURE 3, page 9 of 12

GGTGAGGGGGTACTCCCTCCTCTCCACTCTGCCCTTCCCTCCTGAGAATCCCAGGATGTG  
AGGATGGGAAGAGCTCTTAGCAGCCACCTCACCCATCCATCTTGTAGGACAGAGGCATCC  
TGGGGGTAGGGCAGTAGTGTGGGCAGACTTCCCTCTCCCAGGGATTCCCTCTCTGTTC

10159 GGGGGTGATCAGATCTGTGTGTCCCAGGTGTCTCTGGATCAGTATCAGACGGAGGATGAG  
CTGTACCAGCTGTCCCTGCAGCGGGAGCCGCGCTCCAAGTCTCGGTGAGGGGGTACTCC  
CTCCTCTCCACTCTGCCCTTCCCTCCTGAGAATCCCAGGATGTGAGGATGGGAAGAGCTC  
TTAGCAGCCACCTCACCCATCCATCTTGTAGGACAGAGGCATCCTGGGGGTAGGGCAGTA  
GTGTTGGGCAGACTTCCCTCTCCCAGGGATTCCCTCTCTGTTCCTCCGGGGCTCTGGGCT  
[T,C]  
CCCCTGCCCTCTGGCCCTAGCTCAGGCCCGACCATTTCATAGCCAACCAGCCCCACGAGT  
TGCACCCACCACCCCGGCCCGCGTACTGGAGGAGTGGACCTCGGCTGCCAAACCCAAG  
CTGGATCAGGCCCTCGTGGTGGAGCACATCGAGAAGATGGTGGAGGTGAGCTCCTGCGGA  
GCCTGAGCAGTGTGTGGGGAGAGGCCAGTTTGCCGGAGCACTGCCCTGGAAGCCAGCAG  
AGTGTCTGTTC AAGACCCAGCACTCAGCCCCCTAGGAGTCACAGGGCCTGCGAGGCCAGC

12025 TGGGGTTTTACCATGTTGGCCAGGCTGGTCTCAAACCTCTTGACCTCAAGTGATCTACCTG  
CCTTGGCCCCCAAGTGCTAGGATTACAGGCATGAGCTACTGCTCCTAGCCTAAAAAA  
TTTTTTTTGGGCATGGGTGGCAGTGCCTGTAGTCCCAGCTACTCAGGAGGCTGAGGCAG  
GAGGAACCCCTTGAGCCCAGGAGGTTGAGACTGCAGTGAGCTGTATCACACCACTGCAC  
TCAGCCTGGGTGACTGCGCGAGATCACCCCATCAAAAAAAAAAAAAAAAAAGAAAAAAAA  
[A,-,G]  
GAAGAAATGAAAGTCCCTCTTTCCCTTTCCACTGGTAGAAGTTGCCATGATTAAGCACT  
GTTAACAATATTAAGCTTGGCAGTATGTGGATTCTTCCAGTCTTCTTTTCCCAGGCAGGT  
GCACATTGATAGAGATTTTGTGTTGTTGGTGTCTGTTTCATGGACAAACAGGATTAGAGC  
ATAAATCTAGTTCTGCTTGTGGCTTTTATCATAGCTGCTTTATTTCTTCTCCAGATTTT  
AGGCAGAGGTAGTTGAGTTCCATGTTTTCTCCCTGGGTGGTGGGTGGATTTTTATCTAG

14723 GGCTTCTGGGATTTGTGTTGCGGGGTAGAAAGGCCCTCAGCCCCCAAGATTATAAAATTA  
TAAACCTTTTCTTTTTTTTTTTTTTCTGAGACAGGGTGTCTTGCCATGTCACCCAGG  
CTGGAGTGCAGTGGCATGATCTTGGCTCGCTGCAACCTCCACCTCCCAGGTTCAAGTGAT  
TCTCGTGCCTTAGCCTCCCGAGTAGCTGGGATTATAGGTGCTGCCACTATGCCTGGCTA  
ATTTTTGTATTTTAGTAGAGACGGGGCTTTGCCATGTTGGCCAGGCTGGTCTTGAAC  
[T,C]  
CTGACCTCGTGATCCACCCGCTTGGCCTCCCAAAGTGCTGGGACTACAGGCGTAAGCCA  
CTGTGCTCGGCCCTATATTTTTTTT CAGATAGCCAGTTATCCTAATGCTCCCTTGATTGTA  
TGGACCACCTGGATCACACATTATGAGCCCCCTCATAAGCAGGTGGGAGTCTCAAGCGAG  
GGCCAGTCCCGATGGGAATAGCACTTGGTGGCTGAGGACCCCTCTATCTGTGCAGACACT  
GTTGTAAACTTCACATGCATCATCTAATTTAGTCCTCACCAAATCCTATGAAATGTAG

14996 CCATGTTGGCCAGGCTGGTCTTGAACCTCCTGACCTCGTGATCCACCCGCTTGGCCTCCC  
AAAGTGCTGGGACTACAGGCGTAAGCCACTGTGCTCGGCCCTATATTTTTTTT CAGATAGC  
CAGTTATCCTAATGCTCCCTTGATTTGATGGACCCTGGATCACACATTATGAGCCCCC  
TCATAAGCAGGTGGGAGTCTCAGCGAGGGCCAGTCCCGATGGGAATAGCACTTGGTGGC  
TGAGGACCCCTCCTATCTGTGCAGACACTGTTGTAAACTTCACATGCATCATCTAATTTA  
[G,A]  
TCCTCACCAAATCCTATGAAATGTAGGAATGATCATTACACCCATTTATAGATAAGGAA  
ACGGAGGGACAGGAGATTACTCCGCTACAGGTCAAGAGGCAGGGAAGTAGAGCTGCGAT  
TTGAACTGAGGTCTGTGTCTAGAACACGTGCTCATTCTTTCCCTAAATGTATTATAGG  
TGAAAAAGGGCTTCTGCGGAAAGCCCTGGGTATGTGGGAAACCCTGGATTACAGCTGT  
CTTTCCAGCAGGATGATGCAGGAGAGAGAGGGATGCGATTTTCCCAATCTCTCCTGGTC

16153 CGCCCTCCAGCCCCGGCCCCGCCCTCCCTTCTGGCCCCGCTCTGCCAGAGCCCTTCTC  
AAGCCAGGAAAACCTGGTAATTCTATTTGCTCTCCTCCTGTGTTCTGCCCGGGGCCCT  
GAGGCGGGCTCTAAAGCCCTAGTCTCACCTCAAGAAGGAAGAAGTAGAGTCATCACCTC  
TAAATCCCTCCTCCACACGGCCCCCTCTATTTGCAGATCCTGGGCATCTACAAGCAG  
GGCTCAAATGCCGAGGTGAGATGGAATGACTGGAAGGGCTGCTGGGCAGTGTTTTTTTT  
[T,G]  
TTTGTGTTGTTGTTTGGGAGAGTTACTATTTTGGTGGGGCAATGCCAAGGAGTGAAGTA  
CCTTAAAATCAGAGGCGCATGGCCGGGCATGGTGGCTCAAGCCTGTAATCCCAGCACTTT  
GGGAGGCCGAGGCGCGCAGATCACCTGAGGTGAGGAGTTCAAGACCAGCCTGACCAACAT  
AGCGCAACCCCGCTCTACTAAAAATACAAAAGTAGCTGGGCGTGGTGGCACCCACCTG  
TAATCCCAGCTACTTGGGAGGCTGAGGCATGAGAATCGCTTGAACCTGGGAGGCGGGGT

FIGURE 3, page 10 of 12

16181 CTTCTGGCCCCGCTCTGCCAGAGCCCTTCTCAAGCCAGGAAAACCTGGTAATTCTATTT  
GCCTCTCCTCCTGTGGTTCTGCCCCGGGGCCCTGAGGCGGGCTCTAAAGCCCTAGTCTCAC  
CCTCAAGAAGGAAGAAGTAGAGTCATCACCTCTAAATCCCTCCTCCCACCACGGCCCTC  
CTCTATTGCAGATCCTGGGCATCTACAAGCAGGGCCTCAAATGCCGAGGTGAGATGGAAT  
GACTGGAAGGCTGCTGGGCAGTGTTTTTTTGTTTGTGTTGTTGTTGTTGGGAGAGTTACT  
[G, A]  
TTTTGGTGGGGCAATTGCCAAGGAGTGAAGTACCTTAAAATCAGAGGCGCATGGCCGGGC  
ATGGTGGCTCAAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGCGCAGATCACCTGA  
GGTCAGGAGTTCAAGACCAGCCTGACCAACATAGCGCAACCCCGCCTCTACTAAAAATAC  
AAAAAGTAGCTGGGCGTGGTGGCACCACCTGTAATCCCAGCTACTTGGGAGGCTGAGGC  
ATGAGAATCGCTGAACCTGGGAGGCGGGGTTGCAGTGAGCCGAGATCACGCCACTGCA

16756 CAGTGAGCCGAGATCACGCCACTGCACTCCAGCCTGGGCAACAGAGAGGGCTCTGTCTCA  
AAAAAAAAAAAAACAACAAAAAACCCCAAAACCAAAACCCCAACAAATCAGAGGCTCAAG  
ATGACTGATGTGAAGGGAGTGGCGTTTAAAGAGGCCATTTATTTTGATGACGCAGCTGCCC  
AGGAACAGAGAACATGGGAGAAGGCATAGACTGACAATTAGGAGGAGGAGAACACTTTGG  
AAGGAGACTCTTATTTTGGTGGGGCAGCTGCTCAGGAACAAGGTTCTTGGTAGGGGGGC  
[A, G]  
CAAGCCTGCGGGATGGGATGGAGGGTATTCTGACCAATGTCCCTGGCTGGCTCTCCATTT  
GCTCTCCCCAGCCTGTGGAGTGAAGTGCACAAAGCAGTGCAAGGATCGCCTGTGAGTTG  
AGTGTGCGCGCAGGGCCCAGAGTGTGAGCCTGGAGGGGTCTGCACCCTCACCCCTACCCCA  
TGACAGCCACCATCACCGCGCCTTCACTTCTCTGCCCCGCCCTGGCAGGCGAGGCT  
CCAGGCCTCCAGGTAAGAGGGAGTCATTCTGTACTGGCCTGTGGAGGGAAGGATGCAGGG

18059 AATGATTATTTTGTGAGAACAGTCCGAACACTATGTTAACTGGGGTCTAAGGTAGTT  
GATCACAACCTGTTTGGGTGGCATAAGTCTCTAAAAACAGAGGCAGGCACAGGGCATAAC  
ATCCTCAAAAATAGAAAAGATAAATCCATTTGCATTGAGCCTTCCAGAAGTGTGGGGTC  
TAAATGTGAAATACACACAAAATTGACATTTAAGCAAACCTGCGCTGACAAATCTGTGGC  
TGAAAAAGCTGTGGCAAAACAAAACATAGAAAAGAGCCTCAAAAATTGGGCTGAGGCC  
[A, G]  
GGCATGGTGGCTCACGCCCTGTAATCCTAGCACTTTGGGAAGCCAAGGTGGGTGGATCACC  
CGAGGTGAGGAGTTGGAGACCAGACTGGCCAACGTGGCAAAACATCATCTCTACAATACA  
AAAATACAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGG  
CTGAGGCACGAGAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGAGAGCCGAGATTGCGC  
CATTCGACTCCAGCCTGGGCGACAGAGAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAA

18364 TGGTGGCTCACGCCCTGTAATCCTAGCACTTTGGGAAGCCAAGGTGGGTGGATCACCCGAG  
GTGAGGAGTTGGAGACCAGACTGGCCAACGTGGCAAAACATCATCTCTACAATACAAAAA  
TACAAAATTAGCTGGGCGTGGTGGCAGGCGCCTGTAATCCCAGCTACTTGGGAGGCTGA  
GGCAGCAGAAATCGCTTGAACCTGGGAGGTGGAGGTTGCAGAGAGCCGAGATTGCGCCATT  
GCACTCCAGCCTGGGCGACAGAGAGAGACTCTGTCTCAAAAAAAAAAAAAAAAAAAAAAA  
[A, -, T]  
TGGGCTGTGAGGTGATGCAGGGAATTGATTTTTTGGTGGGTGGGTCTGCTTCTGGGATGAT  
GTGGATGCCTCCCGTGGAGAGGGGAAGGGTTGATGAAGTCCAGGGACCTGGAAGTGTGT  
TCTGCAGCAATCCCCCTCCAGCAGAGATCCGTGAGGAGGAGGTACAGACGGTGGAGGAT  
GGGGTGTGTTGACATCCACTTGAATAGATGGTGAATCCTCCACAGCTGGCACCAGAGCT  
CCCCACTGAGGCTGGGGGGAGCTGGGGAGTATCAGGGAATGGGTGCTTTATCCAAAT

18861 ACTTGTAATAGATGGTGAATCCTCCACAGCTGGCACCAGAGCTCCCCACTGAGGGCTGG  
GGGGAGCTGGGGAGTATCAGGGAATGGGTGCTTTATCCAAATGGCTCCAAGCCAGGTG  
GGCTACTACCTTGTGTTAGGGGGGTGTCTTCTCACAACCTGTTTTTCTTCCAGCT  
GTGGTTGGATCAAGGACTCATTCTGCCTTGGAGAAAATACTTCAACCAGAGCAGGGAGC  
CTGGGGGTGTGGGGCAGGAGGCTGGGGATGGGGGTGGGATATGAGGGTGGCATGCAGCT  
[G, A]  
AGGCAGGGCCAGGGCTGGTGTCCCTAAGGTTGTACAGACTCTTGTGAATATTTGTATTT  
TCCAGATGGAATAAAAAGGCCCGTGAATTAACCTTCACCATCAGCGCCTAGAATCCCGG  
GGGGTAGGGGATGGTATACTTTACAGGATGACAATCTTGGGAGCTAGAACTTTGTAGCC  
AGAGAACTTGGGAGGTCTGAATCTCATGTGTCTGGAGTCTTGGGGAAGAGAATCTTAG  
AAGCAGAAAACCTTGAACATAAGAATCTTGGGGAGGGTCTAGGATCTTGAAGAGACCAG

20443 TGTGTGAAGAGCAGCTCGCTCCTGTGCCGCTGCCTCCTGTGCTGCCTCCATCCCTGCA  
GCCAGTCGGTTCCTCTTGGCTCCTCTCGTCACTACCCCTCCAGTTCAGTCTGGCCTCTT

FIGURE 3, page 11 of 12

CCTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATGCATGCATGCATATGTGTGTCC  
 AGGTCTGCCTGCCCCGGATGTGACAAGTAGCGGTCTTCATGGTTGCATGTGTCTGAATTT  
 GGTGTCTGAGCTTCACATTGTATGCGCCTGTGTGCATGTGTGTGCATGGACATGCATGCT  
 [G,A]  
 TATCTGCTGTGTTTTCCCCTCCCCATGTGTCCCCACTGGCCTTTGCACATGGGAGAAGGG  
 CATGTGCTCAGCATATCACTCAACTGTCCACATTGGGTGGGTACCTGTGTGTGGTGTGTG  
 TGTGTGGGGGTGTGTCTTGAAGTGGCAGGTCCCAAATGCTTAGGCAATCTGAACCTTGG  
 ACCTTGCAGAGAGGAGAGATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGC  
 CCGGTGACCTTTTCTGCCCTTGATGGGCAAAGCTGGGGGTAGGGAAAGGAGACAAGTGCT  
 20881 TTGAAGTGGCAGGTCCCAAATGCTTAGGCAATCTGAACCTTGGACCTTGCAGAGAGGAGA  
 GATGTCCCTGTAGGTGGGAGGGACAGGGAGATGCAGCAGCTGCCCGGTGACCTTTTCTGC  
 CCTTGATGGGCAAAGCTGGGGGTAGGGAAAGGAGACAAGTGCTCATACTTACCTCCCTCC  
 CTGCCCAGGCTCCTCTGTAAAGGTCTGAGTCTGTCTCTGTGAGCCATTGCATCTGTCTGT  
 CTATGCCCTGATGCCTGGATGGACAAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG  
 [A,T]  
 GTGTGAGGCTGCAGGAAGAGGAACAGTGGGGGATGGGCAGGAAAGTGGGCTGTGGGGTCA  
 GGGAGGCGAT